

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 31, 2003, 06:13:31 ; Search time 49.5 Seconds
(without alignments)
6418.680 Million cell updates/sec

Title: US-09-598-982-20

Perfect score: 1458

Sequence: 1 gggcccccgcgagaagaat.....cgtgaagcgcgcgcgtcgt 771

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
	Ygapop 10.0 , Ygapext 0.5
	Fgapop 6.0 , Fgapext 7.0
	Delop 6.0 , Delext 7.0

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xih
-O=/cgr2.1/USPTO.spool/US09598982/runat_27012003_073716_19567/app_query.fasta_1.967
-DB=SPREMBL_21 -OFMT=fastan -SUFFIX=n2p.rspt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi
-LIST=45 -DOCALLIN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09598982 @CGN.1.1.28 @runat_27012003_073716_19567 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-MARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPREMBL_21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.rvivirus:*
16: sp.bacteriaph:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1342	92.0	275	4	Q96R26
					Q96r26 homo sapien

2	1049	71.9	273	11	Q921N4	Q921n4 mus musculus
3	1022	70.1	233	4	Q96R27	Q96r27 homo sapien
4	1019	69.9	273	6	Q9XSM1	Q9xsm1 ovis aries
5	1005	68.9	237	6	Q29464	Q29464 bos taurus
6	988	67.8	242	4	Q96R25	Q96r25 homo sapien
7	699	47.9	166	4	Q9U018	Q9u018 homo sapien
8	651.5	44.7	130	4	Q8TD16	Q8td16 homo sapien
9	651	44.7	277	6	Q8SQ44	Q8sq44 sus scrofa
10	640	43.9	279	11	Q99MS4	Q99ms4 mus musculus
11	589.5	40.4	321	4	Q96R28	Q96r28 homo sapien
12	568	39.0	274	11	Q924N9	Q924n9 mus musculus
13	552	37.9	115	4	Q96L36	Q96l36 homo sapien
14	544	37.3	331	11	Q98L46	Q98l46 mus musculus
15	543	37.2	369	13	Q9PXX7	Q9pxx7 xenopus lae
16	536.5	36.8	310	11	Q91XC4	Q91xc4 mus musculus
17	534.5	36.7	317	13	Q9DGR3	Q9dgr3 xenopus lae
18	531.5	36.5	310	11	Q9QY29	Q9qy29 mus musculus
19	530	36.4	339	11	Q99L44	Q99l44 mus musculus
20	492.5	33.8	297	11	Q88781	Q88781 rattus rat
21	492	33.7	453	11	Q8VDE0	Q8vde0 mus musculus
22	472	32.4	329	13	Q42272	Q42272 xenopus lae
23	469.5	32.2	335	11	Q8VIF2	Q8vif2 mus musculus
24	469	32.2	624	6	Q95ME7	Q95me7 oryctolagus
25	468	32.1	537	4	Q9BVE1	Q9bve1 homo sapien
26	468	32.1	581	4	Q9BVE2	Q9bve2 homo sapien
27	464.5	31.9	799	11	Q9DB10	Q9db10 mus musculus
28	463.5	31.6	624	11	Q91Y47	Q91y47 mus musculus
29	460.5	31.6	806	6	Q9DAT3	Q9dat3 mus musculus
30	454	31.1	806	6	Q18783	Q18783 macropus eu
31	451	30.9	638	11	Q8R0P5	Q8r0p5 mus musculus
32	451	30.9	643	6	Q97506	Q97506 sus scrofa
33	447.5	30.7	247	13	Q9W705	Q9w705 parallachthy
34	446	30.6	282	11	Q9D413	Q9d413 mus musculus
35	446	30.6	322	11	Q920S2	Q920s2 mus musculus
36	440.5	30.2	277	5	Q96899	Q96899 scolopendra
37	437.5	30.0	812	11	Q9R0W3	Q9r0w3 rattus norv
38	437	30.0	1524	13	Q91674	Q91674 xenopus lae
39	434	29.8	267	5	Q98K47	Q98k47 ludia foli
40	434	29.8	334	6	Q46507	Q46507 papio hamad
41	433	29.8	364	5	Q91TV4	Q91tv4 dirosophila
42	430.5	29.5	249	13	Q9W701	Q9w701 parallachthy
43	430	29.5	810	4	Q15146	Q15146 homo sapien
44	429	29.4	435	11	Q8VCA5	Q8vca5 mus musculus
45	429	29.4	454	6	Q46506	Q46506 papio hamad

ALIGNMENTS

RESULT 1

ID	Q96R26	PRELIMINARY:	PRT:	275 AA.
AC	Q96R26			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Trypsinase I.			
GN	TRYPSINASE.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21096910; PubMed=1157797;			
RA	Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,			
RA	Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,			
RA	Higgs D.R.;			
RT	*Sequence, structure and pathology of the fully annotated terminal 2			
RT	Mb of the short arm of human chromosome 16.*;			
RL	Hum. Mol. Genet. 10:339-352(2001).			
DR	EMBL: AB006466; AAK61271.1; ..			
DR	MEROPS: S01.242; ..			
DR	InterPro: IPR001254; Ser_protease_Try.			

DR Pfam: PF00089; trypsin: 1.
 DR PROSITE: PS50240; TRYPsin_DOM: 1.
 DR PROSITE: PS00134; TRYPsin_HIS: UNKNOWN_1.
 DR PROSITE: PS00135; TRYPsin_SER: UNKNOWN_1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 275 AA; 30601 MW; 2F4AAAE7DDE483E CRC64;

Alignment Scores:

Pred. No.:	1,54e-94	Length:	275
Score:	1342.00	Matches:	240
Percent Similarity:	97.96%	Conservative:	0
Best Local Similarity:	97.96%	Mismatches:	5
Query Match:	92.04%	Indels:	0
DB:	4	Gaps:	0

US-09-598-982-20 (1-771) x Q96RZ6 (1-275)

```

OY 19 ATCGTCGGGGGTACAGAGGCCGCCAGAGCAAGTGGCCCTGGAGTGGAGCTGAGAGTC 78
    |||
Db 31 ILevalGIgylgIngluAlaIProArSerIyTrPrProIrgInValSerLeuArGVal 50
OY 79 CAGGAGCCCAATCTGGATGCACTTCTGGGGGGCTCCCTCATCCACCCCGAGTGGTCTG 138
    |||
Db 51 HisAspProTyTrPrMetHisPhcysGlySerLeuIleHisProGIInTrPrValIleu 70
OY 139 ACCGCGGGGGGTGGGAGCGGAGCGGAGTCAAGGATCTGGCCGCCCTCAGGGGTCAACTG 198
    |||
Db 71 ThrAlaAlaHisCysValGlyProAspValIlyAspLeuAlaIleValGInleu 90
OY 199 CGGAGAGCAGCACTCTACTACAGAGCAGCTGCTGCCGGTCAAGAGATCATCTGTCAC 258
    |||
Db 91 ArgIuGInHisLeuTyTrGInAspGIInleuProValSerArGLeIleValHis 110
OY 259 CCACAGTCTTCACCCGCCAATGAGGAGCGACATCGCCCTGGCTGGAGCTGGAGAGCGG 318
    |||
Db 111 ProGIInPrHeTyTrAlaGInIleGIyAlaAspIleAlaLeuGIInleuGIInPro 130
OY 319 GTGAAGGTCTCCAGAGCAGTCCACAGGTCACCTGCCCTCCCTCAGAGCTTCCGCC 378
    |||
Db 131 ValAsnValSerSerHisValHisThrValThrLeuProProIaIaSerGIInThrPhero 150
OY 379 CCGGGGATGCGGTGCTGGTCACTGGCTGGGGCGATGTGAACATGATGAGCGCTCCCA 438
    |||
Db 151 ProGIInMetProCysTyTrPrValThrGIyTrPrGIyAspValAspAsnAspGIInArGIInPro 170
OY 439 CCGGCATTTCTCTGAACAGGTAAGTCCCATTAATGAAGAAACCAATTTGGAGCGCA 498
    |||
Db 171 ProProIleProLeuIlySGInValIlyValProIleMetGIInAsnHisIleCysAspAla 190
OY 499 AAATACACACTTGGGCGCTACACGGAGAGCAGCTCCGATCTGCCCTGAGACATGCTG 558
    |||
Db 191 LysTyTrHisLeuGIInAlaTyTrThrGIyAspAspValArGIInLeValArGIInAspMetLeu 210
OY 559 TGTGGCGGGAACACCCGGAGAGTCAATGCAAGGGCGACTCCGAGGGCCCTGGTGTGC 618
    |||
Db 211 CysAlaGIInAlaThrArGIInArGIInAspSerCysGIInGIyAspSerGIInGIyProLeuValIys 230
OY 619 AAGGTGATGGCACTGCTGAGCGGAGCGGCGTGCACCTGGGGGAGAGGGGTGCCAG 678
    |||
Db 231 LysValAlaSGInTyTrPrLeuGIInAlaGIyValAlaArGIInGIyGIInGIyCysAlaGIIn 250
OY 679 CCCAAGCGGGCTGGAGTATACACCGTGCACCTACTGACTGGAGTGGAGTCCACCATAT 738
    |||
Db 251 ProAsnAlaGIInProGIInIleTyTrThrArGIInValThrTyTrIleAspTrPrIleHis**TyTr 270
OY 739 GTCCCAAAAAAGCG 753
    |||
Db 271 ValProIlyLysPro 275
  
```

RESULT 2

ID 0921N4 PRELIMINARY; PRT: 273 AA.
 AC 0921N4;

DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Mast cell protease 7.
 GN MCP7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Granulata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID-10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBU databases.
 DR EMBL; BC011328; AAH11328.1; -
 DR MEROPS; S01.026; -
 DR MGD; MGI:96943; MCP7.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin: 1.
 DR PROSITE; PS50240; TRYPsin_DOM: 1.
 DR PROSITE; PS00134; TRYPsin_HIS: UNKNOWN_1.
 DR PROSITE; PS00135; TRYPsin_SER: UNKNOWN_1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 273 AA; 30332 MW; 51417476514035BE CRC64;

Alignment Scores:

Pred. No.:	3.82e-72	Length:	273
Score:	1049.00 <td>Matches:</td> <td>188 </td>	Matches:	188
Percent Similarity:	81.67% <td>Conservative:</td> <td>17 </td>	Conservative:	17
Best Local Similarity:	74.90% <td>Mismatches:</td> <td>44 </td>	Mismatches:	44
Query Match:	71.95% <td>Indels:</td> <td>2 </td>	Indels:	2
DB:	11 <td>Gaps:</td> <td>1 </td>	Gaps:	1

US-09-598-982-20 (1-771) x Q921N4 (1-273)

```

OY 1 GGGCCCTCCGAAAGAA-----ATCGTCGGGGGTACAGAGGCCGCCAGAGCAAGTGG 54
    |||
Db 21 GLyProAlaMetThrArGIInGIyGIInIleValGIyGIInGIInAlaHisGIyAsnIlyStr 40
OY 55 CCGTCGAGAGTGAAGCTCTGAGAGTCCAGAGCGGCCCATATCTGATGCACTTCTCGGGGGCTCC 114
    |||
Db 41 ProTrPrGIInValSerLeuArGIInAlaAsnAspThrTyTrPrMetHisPhcysGIInGIySer 60
OY 115 CTCATCCACCCCAAGTGGGTGTGACCGCGCGCGCTGGCTGGAGCGGAGCTCAAGAT 174
    |||
Db 61 LeuIleHisProGIInTrPrValIleuThrAlaAlaHisCysValGIyProAspValAlaAsp 80
OY 175 CTGGCGCCCGCTCAGAGGAGCAAGTCCGGGAGAGCACTTACTACTGACAGCAAGTCTGCTG 234
    |||
Db 81 ProAsnIlyValArGIInGIInleuArGIInIlySGInIlyTyTrIleHisAspHisLeuMet 100
OY 235 CCGGTCAAGAGATCATCTGTCACCCAGTCTTACACCGCCAGATCGAGCGGACATC 294
    |||
Db 101 ThrValSerGIInIleIleThrHisProAspPhcTyTrIleValGIInAspIle 120
OY 295 GCGCTGTGAGCTGTGAGAGACCGGCTGAAGTCTCCAGACCTCCACAGGTTCACCTG 354
    |||
Db 121 AlaLeuLeuIlyStrThrAsnProValAsnIleSerAspTyTrValHisProValProLeu 140
OY 355 CCGCTGCTCAGAGACTTCCCGCGGGAGATGCCGCTGGTGGTGCACCTGGCTGGGGCGAT 414
    |||
Db 141 ProProIaIaSerGIInThrPhcProSerGIyThrIleCysTyTrValThrGIyTrPrGIyAsn 160
OY 415 GTGACATATGAGAGCGCTCCACCGGCATTTCTCTGAGCAGAGTGAAGTCCCATAT 474
    |||
Db 161 IleAspAsnGIyValAsnLeuProProIleProIleProIleProIleProIleProIle 180
OY 475 ATGGAAGAACACATTTGTGAGCGCAAAATACCACTTGGCGGCTCAGAGGAGAGAGAGTGC 534
    |||
Db 181 IleGIInAsnHisLeuCysAspLeuIlyStrHisIlySGIyLeuIleThrGIyAspAsnVal 200
OY 535 CGCATCTGCCGTGAGAGAGCATCTGTCGGGGAACACCGGAGAGAGTCAATGCCAGGCG 594
    |||
Db 201 HisIleValArGIInAspMetLeuIleCysAlaGIyAsnGIInGIyHisAspSerCysGIInGIy 220
  
```

```

OY 595 GACTCGGAGGGCCCCGTGTGTGCAGATGATGACACTGGCTGCAGGGGGCGTGTG 654
Db 221 AspergilyProleuValCysLysValGluAspThrProleuGlnAlaGlyVal 240
OY 655 AGCTGGGGGCGAGGGCTGTGCCCAAGCCAGCCGGCTGGCATCPACCCCTGTACCTAC 714
Db 241 SerTrpLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 260
OY 715 TACTTGACATGATCCACCATGATGATGATGATGATGATGATGATGATGATGATGAT 747
Db 261 TyrLeuAspTrpLysLysLysLysLysLysLysLysLysLysLysLysLysLys 271

RESULT 3
O96R27 PRELIMINARY: PRT: 233 AA.
ID O96R27
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Mast cell tryptase beta III.
GN TRYPTASEB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21096910; PubMed=1157797;
RA Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,
RA Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,
RA Higgs D.R.;
RT "Sequence, structure and pathology of the fully annotated terminal 2
RT Mb of the short arm of human chromosome 16.";
RL Hum. Mol. Genet. 10:339-352(2001).
DR EMBL: AE006466; MAK61270.1; -.
DR MEROPS: S01.027;
DR InterPro: IPR001254; Ser-protease_Try.
DR Pfam: PF00089; trypsin_1.
DR PROSITE: PS50240; TRYPsin_DOM: 1.
DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 233 AA; 25874 MW; 246A96C71B2CE6D0 CRC64;

Alignment Scores:
Pred. No.: 4,31e-70 Length: 233
Score: 1022.00 Matches: 186
Percent Similarity: 97.38% Conservative: 0
Best Local Similarity: 97.38% Mismatches: 5
Query Match: 70.10% Indels: 0
DB: Gaps: 0

US-09-598-982-20 (1-771) x Q96R27 (1-233)
OY 19 ATGCTGGGGGTCAGAGAGGGCCCCAGAGCAAGTGGCTGCAGGTGAGCTGAGAGTC 78
Db 31 TLevAlGlyGlnGlnAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 50
OY 79 CAGGCGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 138
Db 51 ArgAspArgGlyTyrTrpMetHisPheCysGlySerLeuIleHisProGlnTrpValLeu 70
OY 139 ACCGCGCGCGGCTGCGTGGAGCGGAGCTCAAGATCTGGCGGCGCTGAGGTGAGACTG 198
Db 71 ThrAlaAlaHisCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 90
OY 199 GGGAGGAGCGACCTACTACTACAGAGACAGAGCTGCGCGGTGAGAGATCATGCTGCAC 258
Db 91 ArgGlnGlnHisLeuTyrTyrGlnAspGlnLeuProValSerArgIleLeuValHis 110
OY 239 CCACAGTCTACACCGCCAGATCGAGCGGAGCATGCGCGCTGGAGTGGAGGAGCG 318
Db 111 ProGlnPheTyrThrAlaGlnIleLeuValAspIleAlaLeuLeuGlnIlePro 130

```

```

OY 319 GTGAGGTCTCCAGCCAGCTCCACAGCTGACCTGCCCTGACAGACTTCC 378
Db 131 ValAsnValSerSerHisValHisThrValThrLeuProAlaSerGlnThrPhePro 150
OY 379 CGGGAGTCCCGGTGGGTCACTGGCGGCGATGTGACAAATGATGAGCGCTCCCA 438
Db 151 ProGlnMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGlnArgLeuPro 170
OY 439 CGCGCATTTCTCTGAAGCAGGTGAAGTCCCATTAATGAAACACATTTGTGACGCA 498
Db 171 ProProPheProLeuLysGlnValLysValProIleMetGlnAsnHisLecysAspAla 190
OY 499 AATATCCACCTTGGCGCCCTACAGCGAGAGCGACGTCCGATGCTCGTACGACATGCTG 558
Db 191 LysTyrHisLysGlnValAlaLysThrGlyAspAspValArgIleValArgAspAspMetLeu 210
OY 559 TGTGCGGAGACCGCGAGGAGTCAATGCGAG 591
Db 211 CysAlaGlyAsnThrArgArgAspSerCysGln 221

RESULT 4
O9XSM1 PRELIMINARY: PRT: 273 AA.
ID O9XSM1
AC O9XSM1
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Tryptase (EC 3.4.21.59).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Kuminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RX MEDLINE=20308142; PubMed=10848900;
RA Pemberton A.D., McAliese S.M., Huntley J.F., Collie D.D.S.,
RA Scudamore C.L., McEuen A.R., Walls A.F., Miller H.R.P.;
RT "cDNA sequence of two sheep mast cell tryptases and the differential
RT expression of tryptase and sheep mast cell proteinase-1 in lung,
RT dermis and gastrointestinal tract.";
RL Clin. Exp. Allergy 30:818-883(2000).
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPsin FAMILY.
DR EMBL: Y18223; CAB41988.1; -.
DR HSSP: P20231; IAAO.
DR MEROPS: S01.118; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser-protease_Try.
DR Pfam: PF00089; trypsin_1.
DR PRINTS: PR00722; CHYMOTRYPsin.
DR SMART: SM00020; TRYp_SPC; 1.
DR PROSITE: PS50240; TRYPsin_DOM: 1.
DR PROSITE: PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
KW Hydrolase; Serine protease.
FT CHAIN 29 273 TRYPTASE.
SQ SEQUENCE 273 AA; 30156 MW; 88A19DC779053898 CRC64;

Alignment Scores:
Pred. No.: 7,51e-70 Length: 273
Score: 1019.00 Matches: 176
Percent Similarity: 83.67% Conservative: 29
Best Local Similarity: 71.84% Mismatches: 40
Query Match: 69.89% Indels: 0
DB: Gaps: 0

US-09-598-982-20 (1-771) x Q9XSM1 (1-273)
OY 19 ATGCTGGGGGTCAGAGAGGGCCCCAGAGCAAGTGGCTGCAGGTGAGAGTC 78
Db 111 ProGlnPheTyrThrAlaGlnIleLeuValAspIleAlaLeuLeuGlnIlePro 130

```

```

Db 29 ILEIIEGLYGLYISGLIALAProGlySerArgTrpProTrpGlnValSerLeuArgVal 48
QY 79 CAGGGCCCATCTGGAGTCACTTTCGCGGGGCTCCCTCATCCACCCCAAGTGGTCTG 138
Db 49 ArgAspGlnTyrTrpArgHisGlnCysGlySerLeuIleHisProGlnTrpValLeu 68
QY 139 ACCGGGGGGGGTGGTGGGAGCCGAGCAAGGATCGGCGGGCCCTGAGGTTGCACTG 198
Db 69 ThrAlaAlaHisCysIleGlyProGlnLeuGlnIleProSerAspPheArgValGlnLeu 88
QY 199 CGGAGCAGCACTTACTACAGCAGCAGCAGTCTCCGCTGACAGCATCATCTGAC 258
Db 89 ArgGlnGlnHisLeuTyrTyrGlnAspArgLeuProIleSerArgValIleProHis 108
QY 259 CCACAGTCTACACCGCCACAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGAGCCG 318
Db 109 ProHisTyrTyrMetValGlnAsnGlnIleAspIleAlaLeuLeuGlnLeuIlePro 128
QY 319 GTGAGGAGTCTCCAGCCAGCCACAGCGTACCCCTGCGCCCTCAGAGACTTCCCC 378
Db 129 ValSerIleSerCysHisValArgProValIleProIleProIleSerGlnTrpPhePro 148
QY 379 CCGGGAGTCCCTGCTGGGTCACTGCTGGGCGGATGTGGACATGATGAGCGCCCTCCA 438
Db 149 ProGlnSerGlnCysTrpValIleThrGlyTrpGlnValAsnValAspAsnGlyArgProLeuPro 168
QY 439 CCGCATTTCTCTGACAGCAGTGAAGTCCCATTAATGGAACCAACATTTGTACAGCA 498
Db 169 ProProTyrProLeuLysGlnValLysValProIleValGlnAsnSerValLysSptTrp 188
QY 499 AATATGACACCTTGGCGCTACACGGGAGAGCAGTCCGATCGCTGCGTGCAGCATGCTG 558
Db 189 LysTyrHisSerGlnCysLeuSerThrAspTyrSerValProIleValGlnLysAsnLeu 208
QY 559 TGTGCCGGAGACACCCGAGGAGTCTATCCAGGGCGACTCCGAGGGCCCTGCTGTCG 618
Db 209 CysAlaGlnAspGlyLysTrpGlnAspSerCysGlnGlyAspSerCylGlyProLeuValCys 228
QY 619 AAGGTGAATGGACCTCGGTGAGCGGCGGTGTCAGTGGGGGAGGGGTGTGGCCGAG 678
Db 229 LysValAsnGlnTyrTrpLeuGlnAlaGlyValValSerTrpGlyAspLysCysAlaAsn 248
QY 679 CCCAAGCCGCTGGCATCTACACCGCTGTCACTTACTTACTTGGATGGATGCCAGCACTAT 738
Db 249 ProAspTyrProGlnValLysTrpArgIleIleHisSerTyrLeuAspTrpIleHisGlnTyr 268
QY 739 GTCCCAAAAAGCCG 753
Db 269 ValProGlnIlePro 273

RESULT 5
ID 029464 PRELIMINARY: PRT: 237 AA.
AC 029464;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Trypsinase (EC 3.4.21.59) (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN (1)
RP SEQUENCE FROM N.A.
RC TRISSUE=LIVER CAPSULE;
RX MEDLINE=96203914; PubMed=8620861;
RA Pallao M., Gambacurta A., Fiorucci L., Mignogna G., Barra D.,
RA Ascoli F.;
RT "cDNA cloning and primary structure of trypsinase from bovine mast cells
RT and evidence of the expression of bovine pancreatic trypsin inhibitor
RT mRNA in the same cells.";
RL Eur. J. Biochem. 237:100-105(1996).

```

```

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: X94982; CAA64438.1;
DR HSSP: P20231; JNAO.
DR MEROPS: S01.118; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin.1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRYPSIN_SPC.1.
DR PROSITE: PS0240; TRYPSIN_DOM.1.
DR PROSITE: PS00134; TRYPSIN_HIS. UNKNOWN_1.
KW Hydrolase; Serine protease.
FT NON_TER
FT SEQUENCE 1
SO SEQUENCE 237 AA; 26550 MW; EA537A75294EFBA CRC64;

Alignment Scores:
Pred. No.: 8.62e-69 Length: 237
Score: 1005.00 Matches: 178
Percent Similarity: 84.32% Conservative: 21
Best Local Similarity: 75.42% Mismatches: 37
Query Match: 68.93% Indels: 0
Db: 6 Gaps: 0

US-09-598-982-20 (1-771) x 029464 (1-237)
QY 46 AGCAGTGGCCCTGCGCAGTGAAGCTGAGAGTCCAGGCCCATCTGATGATCTTCTGC 105
Db 2 SerGlnTrpProTrpGlnValSerLeuArgValSerArgGlyTyrTrpArgHisHisCys 21
QY 106 GGGGGCTCCCTCATCCACCCCGATGGGTGTGACCGCGCGCGGCTGGTGGAGCCGAG 165
Db 22 GlyLysSerLeuIleHisProGlnTrpValIleThrAlaAlaHisCysValGlyProGln 41
QY 166 GTCAAGATGTGGCCCGCTCAGGCTGCAAGTGGGAGGAGGAGACACTTACATCCAGAGC 225
Db 42 ValHisGlyProSerTyrPheArgValGlnLeuArgGlnHisLeuTyrTyrGlnAsp 61
QY 226 CAGCTGTGCGCGGTGAGCAGATCATCTGTGCACCCAGTTCATACACCCAGTTCGGA 285
Db 62 GlnLeuLeuProIleSerArgIleIleProHisProAsnGlyTyrSerValLysAsnGly 81
QY 286 GCGGACATCGCCCTGTGAGCTGAGGTGAGAGCCGTTGAAGTCTCCAGCAGCTCCACAGC 345
Db 82 AlaAspIleAlaLeuLeuGlnLeuAspLysLeuValAsnIleSerTrpHisValGlnPro 101
QY 346 GTACACCTTCCCGCTGCTCAGAGACCTTCCCGGGGAGTCCGCTGGGTCTACTGAGC 405
Db 102 ValThrLeuProGlnSerGlnTrpPheProProGlyThrGlnCysTrpValThrGly 121
QY 406 TGGGGGATGTGGACATGATGAGCGCTCCAGCCGCGCATTTCTGTGAAGCAGGTGAAG 465
Db 122 TrpGlnValAsnValAspAsnGlyArgArgLeuProProPheProLeuLysGlnValLys 141
QY 466 GTCCCATTAATGAAGAAACCACTTTGTGACGCAAAATACCACTTGGCGCTTACAGCGGA 525
Db 142 ValProValValGlnAsnSerValCysAspArgLysTyrHisSerCylLeuSerThrGly 161
QY 526 GAGGAGTCCGATGCTGCGTGGAGCAGCATCTGTGTGGCGGAGGAGACACCGGAGGACTCA 585
Db 162 AspAsnValProIleValArgGlnAspMetLeuCysAlaGlyAspSerCylArgAsnPhe 181
QY 586 TGCACAGGCGACTCCGAGAGGCGCCCTGTGTGCAAGGTGAATGAGCAGTCTGACAGCGC 645
Db 182 CysGlnGlnAspSerCylGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAla 201
QY 646 GCGGTGTGAGCTGGGCGCAGGGCTGTGCCAGCCCAACCGGCTTGGCATCTACACCCCT 705
Db 202 GlyValValSerTrpGlnAspLysCysAlaLysProAsnArgProGlyIleTyrTrpArg 221
QY 706 GTACCTACTACTTGGATGAGTCCACCACTATGTGCCAAAAGAGCG 753
Db 222 ValThrSerTyrLeuAspTrpIleHisGlnTyrValProGlnIlePro 237

```


RESULT 6
Q96R25 PRELIMINARY; PRT: 242 AA.
ID 096R25.
AC 096R25.
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Putative mast cell MMP-7-like II tryptase (MMP-7-like-2).
GN MMP-7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21096910; PubMed-1157797;
RA Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,
RA Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,
RA Higgs D.R.;
RT "Sequence, structure and pathology of the fully annotated terminal 2
Mb of the short arm of human chromosome 16.";
RL Hum. Mol. Genet. 10:339-352(2001).
DR EMBL; AF006466; AK61272.1; -;
DR InterPro; IPR001254; Ser-protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 242 AA; 26583 MW; 6E72D8AE1DEA2F1 CRC64;

Alignment Scores:
Pred. No.: 1,72e-67 Length: 242
Score: 988.00 Matches: 173
Percent Similarity: 91.71% Conservative: 15
Best Local Similarity: 84.39% Mismatches: 17
Query Match: 67.76% Indels: 0
DB: 4 Gaps: 0

US-09-598-982-20 (1-771) x Q96R25 (1-242)

QY 19 ATCTGCGGGGTCAGAGAGCCCCAGAGCAAGTGCGCCCTGAGAGCTGAGAGTC 78
DB 38 TLevalGlgYlgnGlnLarProArGserLysTrpTrpGlnValSerLeuArGVal 57
QY 79 CAGCGCCCATCTGATGACACTTTCGCGGGGCTCCCTATCCACCCCAAGTGGTGTG 138
DB 58 ArgGlyProTyLTrpMetHisPheCysGlyGlySerLeuLeHisProGlnTrpValLeu 77
QY 139 ACCGCGCGGCTGGTGGAGCGGAGCGCAAGATCGGCCCTCGAGGTGCAACTG 198
DB 78 ThrAlaAlaHisCysValGlnProArPLeuLysAspLeuAlaLeuArGValGlnLeu 97
QY 199 CGGAGACAGACACCTTACTACAGAGACAGCAAGTGTGCGCGTGCAGAGATTCATGTCGAC 258
DB 98 ArgGlnGlnHisLeuTyLTrpGlnAspGlnLeuLeuProValSerArgLLeileValHis 117
QY 259 CCAKAGTTTAAACGSSCCAGATCGAGAGCGAATCGCCCTGTGGAGCTGAGAGCGCG 318
DB 118 ProGlnPheTyLLeileGlnThrGlyAlaAspRLeileAlaLeuLeuGlnGlnPro 137
QY 319 GGAAGGTCTTCAGCAGCAGTCCAGCGGTCAACCTGCGCCCTCGAGAGCTGAGAGCTGCGCC 378
DB 138 ValAsnHisSerHisLeHisThrValThrLeuProGlnAspLeuLThrPhePro 157
QY 379 CCGGGAGATCCCTGTGGTCACTGGCTGGGCGAGTGTGAGCAATGAGAGCGCTCCCA 438
DB 158 ProGlyMePProCysTrpValThrGlyTrpGlyAspValAspAsnAsnValHisLeuPro 177
QY 439 CCGCATTTCTCTGAGAGAGTGAAGGTCCCATATATGAAACCAACCATTTGTGACGCA 498
DB 178 ProProTyLProLeuLysGlnValGlnValProValValGlnAsnHisLeuCysAsnAla 197

QY 499 AATACACACCTTGCGCCCTACACAGGAGAGAGAGTCCGATCGTCGAGACATGCTG 558
DB 198 GluTrpHisThrGlyLeuHisThrGlyHisSerPheGlnIleValArGAspAspMetLeu 217
QY 559 TGTGCGGAGACACCCGAGGAGCATCATGCGAGGGGACTCCGAGAGGCCCTGTGTGC 618
DB 218 CysAlaGlySerGlnAsnHisAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 237
QY 619 AAGGTGATGCGCAC 633
DB 238 LysValAsnGlyThr 242

RESULT 7

Q90U18 PRELIMINARY; PRT: 166 AA.

ID Q90U18.
AC Q90U18.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Mast cell MMP-7-like I protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99121069; PubMed-9920877;
RA Pallao M., Fejzo M.S., Shayesteh L., Blount J.L., Caughey G.H.;
RA "Characterization of genes encoding known and novel human mast cell
RT tryptases on chromosome 16p13.3.";
RL J. Biol. Chem. 274:3355-3362(1999).
CC -! SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AF099147; AAD17861.1; -;
DR HSSP; P20231; IAAO.
DR MEROPS; S01.054; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser-protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYSP-SP; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolase; Serine protease.
FT NON_TER 166
SQ SEQUENCE 166 AA; 18433 MW; A7AF897A6423E7D5 CRC64;

Alignment Scores:
Pred. No.: 1,99e-45 Length: 166
Score: 699.00 Matches: 124
Percent Similarity: 94.85% Conservative: 5
Best Local Similarity: 91.18% Mismatches: 7
Query Match: 47.94% Indels: 0
DB: 4 Gaps: 0

US-09-598-982-20 (1-771) x Q90U18 (1-166)

QY 19 ATCTGCGGGGTCAGAGAGCCCCAGAGCAAGTGCGCCCTGAGAGCTGAGAGTC 78
DB 31 TLevalGlgYlgnGlnLarProArGserLysTrpTrpGlnValSerLeuArGVal 50
QY 79 CAGCGCCCATCTGATGACACTTTCGCGGGGCTCCCTATCCACCCCAAGTGGTGTG 138
DB 51 ArgGlyProTyLTrpMetHisPheCysGlyGlySerLeuLeHisProGlnTrpValLeu 70
QY 139 ACCGCGCGGCTGGTGGAGCGGAGCGCAAGATCGGCCCTCGAGGTGCAACTG 198
DB 71 ThrAlaAlaHisCysMetGlnProArPLeuLysAspLeuAlaLeuArGValGlnLeu 90
QY 199 CGGAGACAGACACCTTACTACAGAGACAGCAAGTGTGCGCGTGCAGAGATTCATGTCGAC 258
DB 91 ArgGlnGlnHisLeuTyLTrpGlnAspGlnLeuLeuProValSerArgLLeileValHis 110


```

Db 207 ALaaspmetleucysAlaGlySerIysGlyLeuAspSerCysGlnGlyAspSerGlyGly 226
Qy 607 CCCCTGCTGTCAGAGTGAATGCACTGGCTGTCAGAGCGCGGTGGTCACTGAGGCGAG 666
    |||||:||||| |||||:|||||:||||| |||||:|||||:|||||
Db 227 ProleuMeCysSerTrpAsnGlyThrTrpValGlnValGlyIleValSerTrpGlyArg 246
Qy 667 GCGTGTGCCACGCCACCGCGCTGGCATCTACACCGCGTGCACCTACTTGACTGG 726
    ||||| ||||| |||||:||||| ||||| |||||:|||||
Db 247 GlyCysGlyLeuHisAsnProGlyValTyrIleArgValMetSerTyrValSerTrp 266
Qy 727 ATCCACCACTATGCTCCCAAAAAGCCG 753
    |||||:|||||:|||||:|||||
Db 267 IleTyrGlnTyrValProArgSerPro 275
    |||||:|||||:|||||:|||||

RESULT 10
Q99MS4 PRELIMINARY: PRT: 279 AA.
AC 099MS4:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Implantation serine proteinase 2.
GN ISP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21361749; PubMed=11467974;
RA O'Sullivan C.M., Liu S.Y., Rancourt S.L., Rancourt D.E.;
RT "Regulation of the trypsin-related proteinase ISP2 by progesterone in
RT endometrial gland epithelium during implantation in mice.";
RN Reproduction 122:235-244(2001).
RM [2]
RE SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=UTERUS;
RA Shen Q.-X., Wang J., Huang Z.-P.;
RT "Identification of endometrial factors involved in mouse embryo
RT implantation.";
RN Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL: AF305425; AAK15264.2; -
DR EMBL: AF442819; AAL38005.1; -
DR HSSP: P00763; IDPO.
DR MEROPS: S01.315; -
DR MGD: MGI:2149952; ISP2.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KM Hydrolase; Serine protease.
SQ SEQUENCE 279 AA; 30986 MW; 81C43A59935244D7 CRC64;

Alignment Scores:
Pred. No.: 7.05e-41
Score: 640.00 Length: 279
Percent Similarity: 64.73% Matches: 120
Best Local Similarity: 49.79% Conservative: 36
Query Match: 43.90% Mismatches: 81
DB: 11 Indels: 4
Gaps: 2

US-09-598-982-20 (1-771) x Q99MS4 (1-279)
Qy 19 ATCGTCGGGCTCAGAGGCGCCGAGAGAGTGGCCCTGCGAGTAGAGCGTAGAGTC 78
    |||||:||||| |||||:|||||:||||| |||||:|||||:|||||
Db 31 IleValGlyGlyHisSerAlaProGlnGlyLysTrpProTrpGlnValSerLeuAlaGlyIle 50
    |||||:|||||:|||||:|||||:||||| |||||:|||||:|||||:|||||
Qy 79 CACGGCCCATAC-----TGATGTCACTTCTGCGGGGGCTCCCTCATCCACCCCGAG 129

```

```

Db 51 TyrArgTyrTyrTrpAlaPheTrpValHisAsnGlyGlySerIleIleHisProGln 70
    :||| |||||:||||| |||||:|||||:||||| |||||:|||||:|||||
Qy 130 TGGGTCTCAGCCGCGCGCGCGTGCCTGGGACCGGACGCTCAAGATGTGGCCCCCAG 189
    |||||:||||| |||||:|||||:||||| |||||:|||||:|||||
Db 71 TrpValLeuThrAlaHisCysIleArgGlnArgAspAlaSerProSerValPheArg 90
Qy 190 GTGCACATCGCGGAGCAGACGCTACTACAGAGACCACTGCTGGCTGGTGCAGCATC 249
    :|||:||||| |||||:|||||:||||| |||||:|||||:|||||
Db 91 IleArgValGlyGlnAlaTyrLeuTyrGlyGlyGlnLeuLeuSerValSerAlaGlyVal 110
Qy 250 ATCGTCACCCACAGTTCTTACACCGCCAGATCGAGCGGACATCGCCCTGTGAGCTG 309
    |||||:||||| |||||:|||||:||||| |||||:|||||:|||||
Db 111 IleIleHisProAspPheValHisAlaGlyLeuGlySerAspValAlaLeuGlnIleu 130
Qy 310 GAGGAGCCGCTGAGAGTCTCCAGCCAGCTCCACAGGTCACCTGCGCCCTGCTAGAG 369
    |||||:||||| |||||:|||||:||||| |||||:|||||:|||||
Db 131 AlaValSerValGlnSerPheProAsnValLysProValLysLeuProSerGlnSerLeu 150
Qy 370 ACCCTCCCGCGCGGCGATGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 429
    |||||:|||||:|||||:|||||:||||| |||||:|||||:|||||:|||||
Db 151 GlnValThrLysLysAspValCysTrpValThrGlyTrpGlyAlaValSerThrHisArg 170
Qy 430 CGCCTCCACCGCCATTTCTCTGAAGCAGAGTGAGAGTCCCATATGAGAAAACACAT 489
    |||||:|||||:|||||:|||||:||||| |||||:|||||:|||||:|||||
Db 171 SerLeuProProTyrArgLeuGlnGlnValGlnValLysIleIleAspSerLeu 190
Qy 490 TGTGACGCCAATATCACCTTGGCGCC---TACAGCGGAGACAGAGCTCCGCATCGTCC 546
    |||||:||||| |||||:|||||:||||| |||||:|||||:|||||
Db 191 CysGlnGlnMetTyrHisAsnAlaThrArgHisArgAsnArgGlyGlnLysLeuIleLeu 210
Qy 547 GACGACATGCTGTCGCGGGAACACCGGAGGAGTCAATGCGAGCGGACATCGCGAGG 606
    |||||:|||||:|||||:|||||:||||| |||||:|||||:|||||:|||||
Db 211 LysAspMetLeuCysAlaGlnValGlnAsnGlnGlyGlnAspSerCysTyrGlyLysPserGlyGly 230
Qy 607 CCCCTGCTGTCAGAGTGAATGCACTGGCTGTCAGAGCGCGGTGGTCACTGAGGCGAG 666
    |||||:||||| |||||:|||||:||||| |||||:|||||:|||||
Db 231 ProleuValLysAsnValThrGlySerTrpThrLeuValGlyValValSerTrpGlyTyr 250
Qy 667 GCGTGTGCCACGCCACCGCGCTGGCATCTACACCGGTGCACCTACTTGACTGG 726
    |||||:||||| |||||:|||||:||||| |||||:|||||:|||||
Db 251 GlyCysAlaLeuArgAspPheProGlyValTyrAlaArgValGlnSerPheLeuProTrp 270
Qy 727 ATC 729
    |||
Db 271 Ile 271

RESULT 11
Q96RZ8 PRELIMINARY: PRT: 321 AA.
ID 096RZ8:
AC 096RZ8:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE HS transmembrane tryptase, gene name TMT, AF175522_1.
GN TRPA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21096910; PubMed=1157797;
RA Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,
RA Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,
RA Higgs D.R.;
RT "Sequence, structure and pathology of the fully annotated terminal 2
RT Mb of the short arm of human chromosome 16.";
RN Hum. Mol. Genet. 10:339-352(2001).
DR EMBL: AE006466; AAK61269.1; -
DR MEROPS: S01.028; -
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00089; trypsin; 1.

```



```

Db 211 Aspmellucysalaglythrserglyarpglyprocyphneglyaspserglyglypro
Oy 610 CTGGTGTGAGAGGTGGACCTGGCTGCAGAGCGGCGGTGCTGAGTGGGGCGAGGGC 669
Db 231 LeuValGlyTrpLysSerAsnLysTrpIleGlnValGlyValSerLysGlyLeasp 250
Oy 670 TGTGGCCAGCCCAACCGGCTGGCATCTACACCCGCTGTCACCTACTACTGTGAGTGCATC 729
Db 251 CysSerAsn---AsnLeuProSerIlePheSerArgValGlnSerSerLeuAlaIrrpIle 269
Oy 730 CACCACTATGTC 741
Db 270 HisGlnHisIle 273

RESULT 13
O96L36 PRELIMINARY; PRT; 115 AA.
AC O96L36;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Delta I tryptase (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hunt J.E., Wang H.W., Thomas P.S., McNeill P.;
RT "Cloning and characterization of novel human trypsinase cDNAs.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY055427; AAL17874.1; -
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin_1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolase; Serine protease.
FT NON_TER 1 1
FT NON_TER 115 115
SQ SEQUENCE 115 AA; 12994 MW; E56C3C592D01BBE9 CRC64;

Alignment Scores:
Pred. No.: 3,23e-34 Length: 115
Score: 552.00 Matches: 100
Percent Similarity: 92.17% Conservative: 6
Best Local Similarity: 86.96% Mismatches: 9
Query Match: 37.86% Indels: 0
Gaps: 0

US-09-598-982-20 (1-771) x O96L36 (1-115)
Oy 112 TCCCTCATCCACCCCATGGGTGCTGACCGCGCGGTGCTGGAGCCGAGCTGAAG 171
Db 1 SerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysMetCluProAspIleLys 20
Oy 172 GATCGCGCGCCCTCAGGCGTCACTGGCGGAGAGCAACCTTACTATACAGAGCAGCTG 231
Db 21 AspleuAlaAlaLeuAlaArgValGlnLeuArgGlnHisLeuTyrGlnIleAspIleLeu 40
Oy 232 CTGCGGTGAGAGGATGATCGTGACCCACAGTTCTACACCGCCAGATCGAGAGCGAGC 291
Db 41 LeuProValSerArgIleIleValHisProGlnPheTyrIleIleGlnThrCluLysAlaAsp 60
Oy 292 ATCGGCTGTGTGAGCTGGAGAGCGCGGTGAAGTCTCAGCAGCAGCTGCACACGGTCAAC 351
Db 61 IleAlaLeuLeuGlnLeuGlnLeuProValAsnIleSerSerHisIleHisThrValThr 80
Oy 352 CTGCCCCCTGCTCAGAGACCTTCCCCCGGAGATGCGCTGGTGCATGCGTGGGGC 411
Db 81 LeuProProAlaSerGluThrPheProProGlnMetProCysTrpValThrIleTyrGly 100
Oy 412 GATGTGACAAATGATGAGCGCTCCACCGCCATTCTCTGAAG 456

```

```

Db 101 AspValAspAsnValHisLeuProSerProTyrProLeuLys 115
RESULT 14
O8R1A6 PRELIMINARY; PRT; 331 AA.
AC O8R1A6;
DT 01-JUN-2002 (TREMblrel. 21, Created)
DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE RIKEN cDNA 2010001P08 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024903; AAH24903.1; -
SQ SEQUENCE 331 AA; 35639 MW; C06F6EF2FA261636 CRC64;

Alignment Scores:
Pred. No.: 1.58e-33 Length: 331
Score: 544.00 Matches: 109
Percent Similarity: 59.30% Conservative: 44
Best Local Similarity: 42.25% Mismatches: 91
Query Match: 37.31% Indels: 14
Gaps: 7

US-09-598-982-20 (1-771) x O8R1A6 (1-331)
Oy 16 AGATCGTGGGGGTGACGAGGAGCCCGCCAGAGCAAGTGCCCTGGCAGTGAGCTGAGA 75
Db 53 ArgIleValSerGlyGlnAspAlaGlnLeuCluArgTyrProTyrGlnValSerValArg 72
Oy 76 GTCCAGCGCCCTACTGATGATGCATCTTGGCGGGGCTCCCTGATCCACCCCGAGGGTG 135
Db 73 GluAsnGly-----AlaHisValCysGlyGlySerLeuIleAlaGluAspIrrpAla 89
Oy 136 CTGACCGCGCGCGGTGCGTGGAGCCGAGCAAGATCTGGCGCCCTCAGGCTGCA 195
Db 90 LeuThrAlaAlaHisCysPheAsnGlnGly---GlnSerLeuSerIleTyrThrValLeu 108
Oy 196 CTGCGGGAGACACCTTACTATACAGGAC-----CAGCTGTGCGCGGTGAGC 243
Db 109 LeuGlyThrIleSerSerTyrProGluAspAsnGluProLysGluLeuArgAlaValAla 128
Oy 244 AGGATCATCGTGCACCCACAGTTCTACACCGCC-----CAGATGGAGCGGACATCGCC 297
Db 123 GlnPheIleLysHisProSer---TyrSerAlaAspGlnHisSerSerIleLysPheAla 147
Oy 298 CTGTGAGAGCTGGAGAGCGCGGTGAAGTCTCCACAGCAAGTGCACAGGTCAACCTGCCC 357
Db 148 LeuValGlnLeuAlaSerProIleSerPheAsnArgTyrMetLeuProValCysLeuPro 167
Oy 358 CTTGCTCAGAACTTCCCGCGGGGATGCGGTGCTGGTGTACTGTGCTGGGGCATGTG 417
Db 168 LysProGlyAspProLeuAspProGlyThrMetCysTyrValThrGlyTyrPheHisIle 187
Oy 418 GACATGATGAGCGGCTCCACCGCCATTTCTGGAAGCAGGTGAAGTCCCAATTAATG 477
Db 188 GlyThrAsnGlnProLeuProProPhePheThrLeuGlnGlnLeuValProLeuIle 207
Oy 478 GAAGAACAATTTGTAGCGCAAAATACACACTTGGCGGCTTACAGCGAGAGACAGCTCCGC 537
Db 208 AspAlaGluThrCysAsnThrTyrTyrGlnGlnAsnSerIleProGlyThrGlu---Pro 226
Oy 538 ATGTCGCGGAGACATGCTGTGTCGGG-----AACACCGGAGAGGACATCAGCCAG 591
Db 227 ValIleLeuGlnGlyMetLeuCysAlaGlyPheGlnGlnGlyLysLysAspAlaCysAsn 246

```

```

QY 592 GCGACCTCGGAGGCCCCGTGTGTGCAAGTGAATGGCACCCTGCGAGCGGCGGTG 651
   |||
Db 247 GYASpSerGlyGlyProLeuValCysAspIleAsnAspValTrrPleInlaGlyVal 266
   |||
QY 652 GTACAGTGGGGGAGGCGGTGTGCCCAAGCCGCGCTGCGATCTACACCCGTGTCAC 711
   |||
Db 267 ValSerTrpGlySerAspCysAlaLeuPheLysArgProGlyValTyrThrAsnValSer 286
   |||
QY 712 TACTACTGTGGATGCACCATATGTCGCCCAAAAGCGCGTGAAGCGGCGCG 765
   |||
Db 287 ValTrrLseTrPleInlaSnThrMetTrrPasnLeuProMetGluGlyArg 304

RESULT 15
Q9PVX7 PRELIMINARY: PRT: 389 AA.
AC Q9PVX7.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Epidermis specific serine protease.
GN XEPSIN.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OX Xenopodinae; Xenopus.
NX NCBI_TaxID=8335;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K.;
RT "The expression control of xepsin by non-axial and planar
RT postteritorizing signals in Xenopus epidermis.";
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
CC -! SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AB018694; BAA84941.1; -
DR HSSP; P00763; IDPO.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 389 AA; 42375 MW; B31FBA42F5D1F6E3 CRC64;

Alignment Scores:
Pred. No.: 1,94e-33 Length: 389
Score: 543.00 Matches: 108
Percent Similarity: 57.14% Conservative: 36
Best Local Similarity: 42.86% Mismatches: 96
Query Match: 37.24% Indels: 12
DB: 13 Gaps: 5

US-09-598-982-20 (1-771) x Q9PVX7 (1-389)
QY 7 CTCGAGAAAGAAATGCTGGGGGCTCAGAGAGGCCCCAGAGCAATGCGCTGAGAGTG 66
   |||
Db 22 IleSerAsnArgIleValGlyGlyMetAspSerLysArgGlyGluTrpProTrpGlnIle 41
   |||
QY 67 AGCGTAGAGTCCACAGCGCCCATACATGATGACATCTCGCGGGGCTCCCTCATCCACCCC 126
   |||
Db 42 SerLeuSerTyrLysSerAsp-----SerIleCysGlyGlySerLeuLeuThrAsp 58
   |||
QY 127 CAGTGGGTGCTGACCGCGCGGCGTGCCTGGAGCGGACGTCAGAGATCTGGCGCCCTC 186
   |||
Db 59 SerTrpValMetThrAlaAlaHisCysIle-----AspSerLeuAspValSerTyrTyr 76
   |||
QY 187 AGGTGCAACTGCGGAGACGACGACTCTACTACGAGACGACGACTGCGCGGTCAAGCAG 246
   |||
Db 77 ThrValTyrLeuGlyAlaTyrGlnLeuSerAlaProAspAsn--SerThrValSerArg 95

```

```

QY 247 -----ATCATGCTGCACCCACAGTCTTCTACACCGCCACGATCGAGCGGACATC 294
   |||
Db 96 GlyValLysSerIleThrLysHisProAspPheGlnTrrGluGlySerSerCysIle 115
   |||
QY 295 GCGCTGCTGGAGCTGGAGAGCGCGGTGAAGGTCTCCAGCCAGCTCCACAGGTCACCTG 354
   |||
Db 116 AlaLeuIleGluLeuGlnLysProValThrPheThrProTrrIleLeuProIleCysLeu 135
   |||
QY 355 CCGCTGCGCTGACAGACATCTCCCGGGGAGCGTGCCTGGGTGACAGGCGGGGCGCAT 414
   |||
Db 136 ProSerGlnAspValGlnPheAlaIleGlyThrMetCysTrpValThrGlyTrpLysn 155
   |||
QY 415 GTGACAATGATGAGCGGCTCCCGCCAGCCATTTCCCTGAGACAGGTGAAGTCCCATAT 474
   |||
Db 156 IleGlnGluGlyThrProLeuLseSerProLysThrIleGlnLysAlaGluValAlaIle 175
   |||
QY 475 ATGAAACACCATTTGTGAGCGCAAAATACACCTTGGCGGCTACAGGAGAGACGATGC 534
   |||
Db 176 IleAspSerSerValCysGlyThrMetTrrGluSerSerLeuGlyTrrIleProAspPhe 195
   |||
QY 535 CGCATCGTCCGTGACGACATGCTGTGCGCGG-----AACACCGGAGGACTCATGC 588
   |||
Db 196 SerPheIleGlnGluAspMetValCysAlaGlyTrrLysGlyGlyArgIleAspAlaCys 215
   |||
QY 589 CAGGCGACTCCGAGAGGCGCCCTGCTGTGCAAGTGAATGGCACCCTGCTGCGAGCGGCG 648
   |||
Db 216 GlnGlyAspSerGlyGlyProLeuValCysAsnValAsnAsnValTrrPleGlnLeuGly 235
   |||
QY 649 GTGTGACGCTGGGAGGCGGTGTGCCCAAGCCCAAGCGGCGCTGATCTACACCCGTGTC 708
   |||
Db 236 IleValSerTrpGlyTrrCylCysAlaGluProAsnArgProGlyValTrrThrLysVal 255
   |||
QY 709 ACCTACTACTGAGTGGATGCACCATATGTCGCC 744
   |||
Db 256 GlnTrrTrrGlnAspTrpLeuLysThrAsnValPro 267

```

Search completed: January 31, 2003, 06:58:26
Job time : 63.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 31, 2003, 05:25:33 : Search time 17.5 Seconds
(without alignments)
3654.657 Million cell updates/sec

Title: US-09-598-982-20

Perfect score: 1458

Sequence: 1 ggagccctcgagaaagaat.....cgtgaagcggcgcgcgcgt 771

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p_model -DEV=xlh
-Q=/cgn2_1/USPRO/spool/US09598982/unal_27012003_073715_19534/app_query.fasta.1.967
-DB=SwissProt.40 -QFMT=fastan -SUFFIX=n2p.rsp -MINMATCH=0.1 -LOOFCI=0
-LOOEXT=0 -UNITS=bits -START=1 -MATRIX=blsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09598982 @CGN.1.1.6 @runat.27012003_073715_19534 -NCP=6 -ICPU=3
-NO_XLPRY -NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-MARKTIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt.40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1368	93.8	275	1	TRB2_HUMAN
2	1363	93.5	275	1	TRB1_HUMAN
3	1282	87.9	275	1	TRXA_HUMAN
4	1092	74.9	276	1	MCT6_MOUSE
5	1077	73.9	275	1	TRYT_CANFA
6	1075	73.7	270	1	TRYT_MERUN
7	1057	72.5	273	1	MCT7_MOUSE
8	1055	72.4	274	1	MCT6_RAT
9	1039	71.3	273	1	MCT7_RAT
10	1030	70.6	273	1	TRYT_SHEEP
11	1006	69.0	275	1	TRYT_PIG
12	988	67.8	235	1	TRXD_HUMAN
13	684.5	46.9	1269	1	TRYM_CANFA
14	608.5	41.7	311	1	TRYG_MOUSE
15	593.5	40.7	321	1	TRYG_HUMAN
16	563.5	38.6	290	1	MPN_HUMAN
17	529	36.3	342	1	PSS8_RAT
18	526	36.1	342	1	PSS8_MOUSE

19	512	35.1	343	1	PSS8_HUMAN	O16651 homo sapien
20	501.5	34.4	317	1	BSS4_HUMAN	O99244 homo sapien
21	495	34.0	455	1	TWS5_MOUSE	Q9er04 mus musculu
22	494	33.9	314	1	TEST_HUMAN	Q9y6m0 mus musculu
23	476.5	32.7	324	1	TEST_MOUSE	Q9jh17 mus musculu
24	474.5	32.5	454	1	TWS3_HUMAN	P57727 mus musculu
25	467	32.0	625	1	FALL_HUMAN	P03951 homo sapien
26	464	31.8	457	1	TMS5_HUMAN	O91353 homo sapien
27	463	31.8	306	1	BSS4_MOUSE	Q9er10 mus musculu
28	460.5	31.6	618	1	HATT_HUMAN	O60235 homo sapien
29	451	30.9	638	1	KAL_HUMAN	P03952 homo sapien
30	448	30.7	338	1	PLMN_HORSE	P06010 equus cabal
31	447	30.7	638	1	KAL_MOUSE	P26521 mus musculu
32	446.5	30.6	812	1	PLMN_BOVIN	P06866 bos taurus
33	445	30.5	638	1	KAL_RAT	P14272 ratius norv
34	444.5	30.5	271	1	EL2_RAT	P00774 ratius norv
35	444	30.5	333	1	PLMN_CANFA	P80009 canis famil
36	443	30.4	245	1	CTRB_BOVIN	P00747 bos taurus
37	441	30.2	343	1	PLMN_SHEEP	P81286 ovis aries
38	439.5	30.1	269	1	EL2A_HUMAN	P08217 homo sapien
39	434	29.8	490	1	TMS2_MOUSE	Q91498 mus musculu
40	430.5	29.5	271	1	EL2_MOUSE	P05208 mus musculu
41	429	29.4	810	1	PLMN_HUMAN	P00747 homo sapien
42	428	29.4	810	1	PLMN_MACMU	P12545 macaca mula
43	427	29.3	263	1	CTRB_CANFA	P04813 canis famil
44	427	29.3	266	1	EL1_PIG	P00772 sus scrofa
45	426	29.2	263	1	CTRB_HUMAN	P17536 homo sapien

ALIGNMENTS

RESULT 1

ID	TRB2_HUMAN	STANDARD:	PRT:	275 AA.
AC	P20231; Q9UQ17; Q9UQ16; Q15664; Q95827;			
DT	01-FEB-1991 (Rel. 17, Created)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Trypsin beta-2 precursor (EC 3.4.21.59) (Trypsin II).			
GN	TPS2 OR TPS2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (VARIANT BETA-2).			
RC	TISSUE=Lung;			
RX	MEDLINE=90369005; PubMed=2203827;			
RA	Miller J.S., Moxley G., Schwartz L.B.;			
RT	"Cloning and characterization of a second complementary DNA for human			
RT	trypsinase.";			
RL	J. Clin. Invest. 86:864-870(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A. (VARIANT BETA-2).			
RX	MEDLINE=93166209; PubMed=8434231;			
RA	Blom T., Hellman L.;			
RT	"Characterization of a trypsinase mRNA expressed in the human basophil			
RT	cell line K0812.";			
RL	Scand. J. Immunol. 37:203-208(1993).			
RN	[3]			
RP	SEQUENCE FROM N.A. (VARIANTS BETA-2 AND BETA-3).			
RX	MEDLINE=90251647; PubMed=2187193;			
RA	Vanderstille P., Ballinger S.M., Tam E.K., Goldstein S.M., Craik C.S.,			
RT	Caughy G.H.;			
RT	"Human mast cell trypsinase: multiple cDNAs and genes reveal a			
RL	multigene serine protease family.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 87:3811-3815(1990).			
RP	[4]			
RP	SEQUENCE FROM N.A. (VARIANTS BETA-2 AND BETA-3).			
RX	MEDLINE=99121069; PubMed=9920877;			
RA	Pallaro M., Fejzo M.S., Shayesteh L., Blount J.L., Caughy G.H.;			
RT	"Characterization of genes encoding known and novel human mast cell			
RT	trypsinases on chromosome 16p13.3.";			

RU	J. Biol. Chem.	274:3355-3362(1999).
RU	[5]	
RU	REVISTONS.	
RA	Palladio M., Fejzo M.S., Shayesteh L., Blount J.L., Caughey G.H.;	
RL	Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.	
RN	[6]	
RN	X-RAY CRYSTALLOGRAPHY (3 0 ANGSTROMS).	
RX	MEDLINE=98180625; PubMed=9521329;	
RA	Pereira P.J.B., Bergner A., Macedo-Ribeiro S., Huber R.,	
RA	Matschner G., Fritz H., Sommerhoff C.P., Bode W.;	
RT	"Human beta-trypsinase is a ring-like tetramer with active sites facing	
RT	a central pore."	
RL	Nature 392:306-311(1998).	
RN	[7]	
RN	X-RAY CRYSTALLOGRAPHY (3.4 ANGSTROMS).	
RX	MEDLINE=99432168; PubMed=10500112;	
RA	Sommerhoff C.P., Bode W., Pereira P.J.B., Stubbs M.T.;	
RA	Stuerzbecher J., Piechotka G.P., Matschner G., Bergner A.;	
RT	"The structure of the human beta-trypsinase tetramer: fo(u)r better or	
RT	worse."	
RL	Proc. Natl. Acad. Sci. U.S.A. 96:10984-10991(1999).	
CC	-I- FUNCTION: TRYPTASE IS THE MAJOR NEUTRAL PROTEASE PRESENT IN MAST	
CC	CELLS AND IS SECRETED UPON THE COUPLED ACTIVATION-DEGRANULATION	
CC	RESPONSE OF THIS CELL TYPE.	
CC	-I- CATALYTIC ACTIVITY: Preferential cleavage: Arg-I-, Lys-I-, but	
CC	with more restricted specificity than trypsin.	
CC	-I- SUBUNIT: HOMOTETRAMER.	
CC	-I- SUBCELLULAR LOCATION: RELEASED FROM THE SECRETORY GRANULES UPON	
CC	MAST CELL ACTIVATION.	
CC	-I- POLYMORPHISM: There are two alleles; beta-II and beta-III which	
CC	differ by 3 residues.	
CC	-I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. TRYPTASE SUBFAMILY.	
CC		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (see http://www.isb-sib.ch/announcement/).	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL: M37482; AAA51843.1; -	
DR	EMBL: M33492; AAA56779.1; -	
DR	EMBL: M33493; AAA36780.1; -	
DR	EMBL: S53551; AAD13876.1; -	
DR	EMBL: AF099143; AAD17859.2; -	
DR	EMBL: AF099145; AAD17857.1; -	
DR	EMBL: AF099146; AAD17858.1; -	
DR	PIR: A37193; A37193.	
DR	PIR: B35863; B35863.	
DR	PDB: 1AAO; 31-JAN-94.	
DR	PDB: 1AOL; 23-MAR-99.	
DR	MEROPS: S01.027; -	
DR	MEROPS: S01.242; -	
DR	Genew, HGNC:14120; TPMSB2.	
DR	InterPro: IPR001314; Chymotrypsin.	
DR	InterPro: IPR001254; Ser_protease-Try.	
DR	Pfam: PF00089; trypsin; 1.	
DR	PRINTS: PR00722; CHYMOTRYPSIN.	
DR	SMART: SMO0020; Tryp_Spc; 1.	
DR	PROSITE: PS50240; TRYPsin_DOM; 1.	
DR	PROSITE: PS00134; TRYPsin_HIS; 1.	
DR	PROSITE: PS00135; TRYPsin_SER; 1.	
KW	Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen;	
KW	Polymorphism; 3D-structure.	
FT	SIGNAL	1 18
FT	PROPEP	19 30
FT	CHAIN	31 275
FT	ACT_SITE	74 74
FT	ACT_SITE	121 121
FT	ACT_SITE	224 224
FT	DISULFID	59 75
FT	DISULFID	155 230

[illegible]

AC Q15661; Q15663; Q9H2Y4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Trypsin beta-1 precursor (EC 3.4.21.59) (Trypsin 1) (Trypsin I).
 GN TPST1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90251647; PubMed=2187193;
 RA Vanderslice P., Ballinger S.M., Tam E.K., Goldstein S.M., Craik C.S.,
 RA Caughey G.H.;
 RT "Human mast cell tryptase: multiple cDNAs and genes reveal a multigene
 RT serine protease family.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3811-3815(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99121069; PubMed=9920877;
 RA Pallaro M., Feljo M.S., Shayesteh L., Blount J.L., Caughey G.H.;
 RT "Characterization of genes encoding known and novel human mast cell
 RT tryptases on chromosome 16p13.3.";
 RL J. Biol. Chem. 274:3355-3362(1999).
 RN [3]
 RP SEQUENCE OF 54-275 FROM N.A. (ISOFORM 2).
 RA Wang H.W., McNeil H.P., Thomas P.S., Murphy B.N., Webster M.J.,
 RA Hellarich A., King G., Heywood G.J., Huang C., Stevens R.L.,
 RA Hunt J.E.;
 RT "Molecular cloning and characterization of novel human tryptase cDNAs
 RT and splicing variants.";
 RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: TRYPTASE IS THE MAJOR NEUTRAL PROTEASE PRESENT IN MAST
 CC CELLS AND IS SECRETED UPON THE COUPLED ACTIVATION-DEGRADATION
 CC RESPONSE OF THIS CELL TYPE.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-I-, Lys-I-, but
 CC with more restricted specificity than trypsin.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: RELEASED FROM THE SECRETORY GRANULES UPON
 CC MAST CELL ACTIVATION.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
 CC produced by alternative splicing.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. TRYPTASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M33494; AAC8172.1; -;
 DR EMBL: M33491; AA36778.1; -;
 DR EMBL: AF099144; AAD17860.1; -;
 DR EMBL: AF206667; AAG35697.1; -;
 DR HSSP: P20231; 1A0L.
 DR MEROPS: S01.242; -;
 DR Genew: HGNC:12019; TPST1.
 DR MIM: 191081; -;
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser_Protease_Try.
 DR Pfam: PF00089; trypsin.1.
 DR PRINTS: PR00722; CHYMOTRYPIN.
 DR SMART: SM00020; Tryp_Spc.1.
 DR PROSITE: PS02440; TRYPSIN_DOM.1.
 DR PROSITE: PS00134; TRYPSIN_HIS.1.
 DR PROSITE: PS00135; TRYPSIN_SER.1.
 KW Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen;
 FT SIGNAL 1 18
 FT PROPEP 19 30 POTENTIAL.
 FT ACTIVATION PEPTIDE (BY SIMILARITY).

FT CHAIN 31 275 TRYPTASE BETA-1.
 FT ACT_SITE 74 74 CHARGE RELAY SYSTEM.
 FT ACT_SITE 121 121 CHARGE RELAY SYSTEM.
 FT ACT_SITE 224 224 CHARGE RELAY SYSTEM.
 FT DISULFID 59 75 BY SIMILARITY.
 FT DISULFID 155 230 BY SIMILARITY.
 FT DISULFID 188 211 BY SIMILARITY.
 FT DISULFID 220 248 BY SIMILARITY.
 FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPALD 79 87 MISSING (IN ISOFORM 2).
 SQ SEQUENCE 275 AA; 30515 MW; ADC48FDC51F37112 CRC64;
 Alignment Scores:
 Pred. No.: 1.38e-72 Length: 275
 Score: 1363.00 Matches: 243
 Percent Similarity: 99.18% Conservative: 0
 Best Local Similarity: 99.18% Mismatches: 2
 Query Match: 93.48% Indels: 0
 DB: Gaps: 0
 US-09-598-982-20 (1-771) x TRB1_HUMAN (1-275)
 QY 19 ATCTGTGGGGCTCAGAGAGGCCCCAGAGACAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 78
 DB 31 IIEVALGIGLYGLNGLIALAARAGSERLYSTRPROTPRGLVALSERLEUARGVAL 50
 QY 79 CAGGCGCCATCTGAGATGACATCTTGGCGGGGCTCCATCCAGCCCGAGTGGGTGCTG 138
 DB 51 HSLGIPROTLYTRPMETHISPhECYSGLYSERLEULENHSIPROGLINTRPALLEU 70
 QY 139 ACCGGCGGCGCTGGGTGGAGCCGACGTCAGAGATCTGGCCGCTCAGGGTGCACACTG 198
 DB 71 THRALAALAHISCYVALGILPRASPVALLYSAPRLEUALAALAVALGVALINLEU 90
 QY 199 CCGGAGCAGACCTCTACTACAGAGCAGACGTCGCGCGGTGAGAGATCTGTCGAC 258
 DB 91 ARGGLUGLNHISLEUPLYTYRGLINSPRGINLEUPEPROVALSERARGILELEVALHIS 110
 QY 239 CCACAGTCTACACCGCCGACATCGGAGCGGACATCGCCCTGTGGAGCTGGAGAGCGG 318
 DB 111 PROGINPHEPLYTHRALAGLNLGGLVALASPRILEALALEULEUGLNUGLNU 130
 QY 319 GTGAGAGTCTCCAGGACGTCACAGGCGGTCACCGCCCGCTCGAGAGCTTCCG 378
 DB 131 VALASVALSERSENHISVALHISTHRVALTHRLEUPROPRVALASERGLTHRPHRO 150
 QY 379 CCGGGATGCGCTGCTGGGTCACTGGCTGGGGCGATGTGGAATGATGAGCGCTCCCA 438
 DB 151 PROGLYMEUPROCYSTRVALTHRGILYTRPGIYASPVALLASPRASPRGLUARGLEU 170
 QY 439 CCGCATTTCTCTGTAAGCAGGTGAAGTCCCATTAATGAAMAACACATTTGTGACCA 498
 DB 171 PROPRPHEPROLEULYGLNVALYVALPROILEMETGLUASNHISILECYASPRALA 190
 QY 499 AATACACACCTTGGGCGCTACACGGGAGAGCGACGTCGCGATGTCGGAGACATCTGCTG 558
 DB 191 LYSTYRHSLEUGLYALATYRTHRGLYASPRVALARGILEVALARGASPRMELEU 210
 QY 559 TGTGGCGGGAACACCGGAGGAGACTCATGCGAGGCGCATCGGAGGCGCCCTGTGTCG 618
 DB 211 CYSALAGLYASNTHRARGASPRSERCYSGNLGLYASPRSEGLGLYPROLEUVALYS 230
 QY 619 AAGTGATGTCGACCTGGCTGAGCGGGCGTGTGCTGAGTGGGGGAGAGGCTGTGCCAG 678
 DB 231 LYSVALASNGLYTHRTRPLEUGLNALGLYVALASERTPRGLYGLNGLYCYSLALAIN 250
 QY 679 CCACACCGCGCTGGCATGTACACCGGTCACCTACTCTGGATGGATGATCCACCATAT 738
 DB 251 PROASNAIYPROGLYIILETYRTHARGVALTHRYTYRLEUASPTPRILEHISHTYR 270
 QY 739 GTCCCGAANAAGCCG 753
 |||||||||||||||

```

Db 271 ValProLysLysPro 275

RESULT 3
ID TRYA_HUMAN STANDARD: PRT; 275 AA.
AC P15157; O90Q11; O9H2Y5.
DT 01-APR-1990 (Rel. 14, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-tryptase precursor (EC 3.4.21.59) (Tryptase 1).
GN Tpsi1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Lung;
RX MEDLINE=90009311; PubMed=2677049;
RA Miller J.S., Westlin E.H., Schwartz L.B.;
RT "Cloning and characterization of complementary DNA for human
RT tryptase.";
RT J. Clin. Invest. 84:1188-1195(1989).
RN [2]
RP REVISIONS TO 89-93 AND 108.
RA Schwartz L.B.;
RN [3]
RP SUBMITTED (MAR-1990) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=99121069; PubMed=9920877;
RA Pallaro M., Fejzo M.S., Shaysien L., Blount J.L., Caughey G.H.;
RT "Characterization of genes encoding known and novel human mast cell
RT tryptases on chromosome 16p13.3.";
RT J. Biol. Chem. 274:3355-3362(1999).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Lung;
RA Wang H.W., McNeil H.P., Thomas P.S., Murphy B.N., Webster M.J.,
RA Heliataratchi A., King G., Heywood G.J., Huang C., Stevens R.L.,
RA Hunt J.E.;
RT "Molecular cloning and characterization of novel human tryptase cDNAs
RT and splicing variants.";
RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 31-50. AND PITUITARY, SEQUENCE OF 31-38.
RC TISSUE=Lung;
RX MEDLINE=87109258; PubMed=3543004;
RA Cromlish J.A., Seidah N.G., Marcinkiewicz M., Hamelin J., Johnson D.A.,
RA Chretien M.;
RT "Human pituitary tryptase: molecular forms, NH2-terminal sequence,
RT immunocytochemical localization, and specificity with prothromone and
RT immunogenic substrates.";
RL J. Biol. Chem. 262:1363-1373(1987).
CC -I- FUNCTION: TRYPTASE IS THE MAJOR NEUTRAL PROTEASE PRESENT IN MAST
CC CELLS AND IS SECRETED UPON THE COUPLED ACTIVATION-DEGRADULATION
CC RESPONSE OF THIS CELL TYPE.
CC -I- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|- , Lys-|- , but
CC with more restricted specificity than trypsin.
CC -I- SUBUNIT: HOMOTETRAMER.
CC -I- SUBCELLULAR LOCATION: RELEASED FROM THE SECRETORY GRANULES UPON
CC MAST CELL ACTIVATION.
CC -I- ALTERNATIVE PRODUCTS: 2 isoforms: 1 (shown here) and 2; are
CC produced by alternative splicing.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. TRYPTASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----

```

DR	EMBL: M30038;	AAAB6934.1;	-	-
DR	EMBL: AF096328;	AAAD1846.1;	-	-
DR	EMBL: AF206665;	AAG3669.1;	-	-
DR	EMBL: AF206666;	AAG3696.1;	-	-
DR	PIR: A45754;	A45754.		
DR	HSSP: P20231;	1AOL.		
DR	MEROPS: S01.015;	-		
DR	MEROPS: S01.143;	-		
DR	Genew: HNC:12018;	TPSL.		
DR	MIM: 191080;	-		
DR	Interpro: IPR001314;	Chymotrypsin.		
DR	Interpro: IPR001254;	Ser. protease_Try.		
DR	Pfam: PF00089;	trypsin. 1.		
DR	PRINTS: PR00722;	CHYMOTRYPSIN.		
DR	SMART: SM00020;	Tryp.Spec. 1.		
DR	PROSITE: PS05240;	TRYPSIN_DOM. 1.		
DR	PROSITE: PS00134;	TRYPSIN_HIS. 1.		
DR	PROSITE: PS00135;	TRYPSIN_SER. 1.		
KM	Hydrolase: Serine protease:	Signal: Glycoprotein; Zymogen:		
KM	Polymorphism: Alternative splicing.			
FT	SIGNAL	1	18	POTENTIAL.
FT	PROPEP	19	30	ACTIVATION PEPTIDE.
FT	CHAIN	31	275	ALPHA-TRYPTASE.
FT	ACT_SITE	74	74	CHARGE RELAY SYSTEM (BY SIMILARITY)
FT	ACT_SITE	121	121	CHARGE RELAY SYSTEM (BY SIMILARITY)
FT	ACT_SITE	224	224	CHARGE RELAY SYSTEM (BY SIMILARITY)
FT	DISULFID	59	75	BY SIMILARITY.
FT	DISULFID	155	230	BY SIMILARITY.
FT	DISULFID	188	211	BY SIMILARITY.
FT	DISULFID	220	248	BY SIMILARITY.
FT	CARBOHYD	132	132	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	233	233	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VASAPPLIC	79	87	MISSING (IN ISOFORM 2).
FT	VARIANT	15	15	R -> P (IN ALPHA-II).
FT				/FTID=VAR_012102.
FT	VARIANT	221	221	K -> Q (IN ALPHA-II).
FT				/FTID=VAR_012103.
FT	CONFILANT	215	216	TR -> SQ (IN REF. 1).
SO	SEQUENCE	275 AA:	30772 MW;	B9BAC4BBC91CE75 C6C64;

Alignment Scores:	
Pred. No.:	7,176-68
Score:	1282.00
Percent Similarity:	94.69%
Best Local Similarity:	93.06%
Query Match:	87.93%
DB:	1
US-09-598-982-20 (1-771) x TRYA_HUMAN (1-275)	
QY	19 ATGCTGGGGGCTCAGAGAGCCCCCAGAGACCAAGTGGCCCTGGCAGGTGAGCTGAGAGATC 78
DB	31 TLevalGlyGlyGlnGlnLaIaProAdgSerLysTTPProTbGlnValSerLeuAryVal 50
QY	79 CACGGCCCTTACCGATGGATGTCAGTCTCGGGGGGCTCCCTATCCACCCAGTCGGTGTG 138
DB	51 ArgAspArgTyrTrpMetHisPheCysGlyGlySerLeuLeuHisProGlnTrpValLeu 70
QY	139 ACCGCGCGCGGCTGCGTGGAGCCGAGCTCAGAGATCTGGCCGCTCAGGTGCACTG 198
DB	71 ThrAlaAlaHisCysLeuGlyProAdgPValLysAspLeuAlaThrLeuAryValGlnLeu 90
QY	199 CGGAGACACACCTCTACACACGAGACGAGCTCGCGGTGCGAGGATCATCGTGCAC 258
DB	91 ArgGlnGlnHisLeuTyrTrpGlnAspGlnLeuLeuProValSerArgLLeuValHis 110
QY	259 CCACACTTGTACACCCGCCAGATCGAGCGGAGACATGCGCCCTGCTGAGGTGAGAGACCG 318
DB	111 ProGlnPheTyrLLeuLeuGlnTrnGlyAlaAspLLeuLeuGlnLeuGlnGluPro 130
QY	319 GTGAAAGCTCCAGGCCAGTCCACAGGATCACTCGCCCTCGCTCCAGAGACTTCCCC 378
DB	131 ValAsnIleSerSerTrpArgValHisTrnValMetLeuProProAlaSerGluTrnPhePro 150

QY	379	CCGGGGAATGCCGCTCCTGGGTACATGCTGGGGGGATGTTGGACAATATGAGCCGCTCCCA	438
Db	151	ProtiGeneProcStrPrvalThrlg1YtP6LyAspAlaSpAsnSpArg1PurProLeuP170	170
QY	439	CCGGCATTTCTCTCTGAAGCAGAGGTGAAGGTCCCATTAATGGAACCAACCATTTGTGACGCA	498
Db	171	ProPhorProleauLysGlnValLysValProIleMetLusnHstIecYsAspAla	190
QY	499	AAATATCACCTTTGGCGCCTTACACGGGAGACAGCTCCGATCGTCCGTGACGACATGCTG	558
Db	191	LysTyrlHisLeuGlnYlaIaTyrlThrlGlyAspAspValArgIleIleIaArgAspMetLeu	210
QY	559	TGTGCGGGGAACACCCGGAGGAGTACTATGCGAGGGGCGACATCCGGAGGGCCCGCTGTGC	618
Db	211	CysAlaGlnsnTrlAlaArgAspSerCysLysGlyAspSerGlyGlyProLeuValCys	230
QY	619	AAGTGTGAATGGCACCTGCTGCACGGCGGGCTGGTGCAGTGGGGCGAGGAGCTGTGCCCA	678
Db	231	LysValAsnGlyTrlTrPleuGlnAlaGlyValValSerTrlPAspArgLugLysAlaGln	250
QY	679	CCCAACGGGCTGGCATCTACACCCGCTGTACCTACTTGGACTGGATGCACCACTAT	738
Db	251	ProAsnArgProGlyIleTyrlThrlArgValThrlTyrlLeuAspTrlIleHisTyr	270
QY	739	GTCCCAAAAGCCG	753
Db	271	ValProLysLysPro	275
RESULT 4			
ID	MCT6_MOUSE	STANDARD;	PRT; 276 AA.
AC	P21845: 061962:		
DT	01-MAY-1991 (Rel. 18, Created)		
DT	01-AUG-1991 (Rel. 19, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Mast cell protease 6 precursor (EC 3.4.21.59) (M MCP-6) (Trypsinase).		
GN	MCP6.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=91139682; PubMed=1995638;		
RA	Reynolds D.S., Gurley D.S., Austen K.F., Serafin W.E.:		
RT	"Cloning of the cDNA and gene of mouse mast cell protease-6.		
RT	Transcription by progenitor mast cells and mast cells of the		
RT	connective tissue subclass.";		
RL	J. Biol. Chem. 266:3847-3853(1991).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Leaden X AL;		
RC	MEDLINE=94023807; PubMed=821098;		
RX	Huang R., Abrink M., Gohl A.E., Nilsson G., Aveskog M., Larsson L.G.		
RA	Nilsson R., Hellman L.:		
RT	"Expression of a mast cell tryptase in the human monocytic cell lines		
RT	U-937 and Mono Mac 6.";		
RL	Scand. J. Immunol. 38:359-367(1993).		
RN	[3]		
RP	SEQUENCE FROM N.A. (SHORT ISOFORM).		
RC	STRAIN=Leaden X AL;		
RX	MEDLINE=95048582; PubMed=7959952;		
RA	Huang R., Hellman L.T.;		
RT	"Genes for mast-cell serine protease and their molecular evolution.";		
RT	Immunogenetics 40:397-414(1994).		
RN	[4]		
RP	SEQUENCE OF 32-54.		
RX	MEDLINE=90222202; PubMed=2326280;		
RA	Reynolds D.S., Stevens R.L., Lane W.S., Carr M.H., Austen K.F.,		
RA	Serafin W.E.:		
RT	"Different mouse mast cell populations express various combinations		
RT	of at least six distinct mast cell serine proteases.";		

Seq	Alignment Scores:	Length:	Matches:	Conservative:	Mismatches:	Indels:
US-09-598-982-20 (1-771) x MCT6_MOUSE (1-276)	8.21e-57	276	192	20	36	2
DB: 1	74.90%					1
0Y 4 CCCCCTGAGAAAGA-ATCGTCGGGGGTCAGAGAGCCCCCAGGAGCAGTGGCCC	Score: 1092.00	276	192	20	36	2
Db 25 ProAlaLaengInaGcValcIYLlEValdIlgYlnhSgdlAlaSerGlUsertYttrPro	Percent Similarity: 84.80%					
0Y 58 TGGCAGGTGACCGTGAAGAGTCCAGGCCCCAATACATGAGCATCTCTGCGGGGCTCCCT	Best local Similarity: 76.80%					
Db 45 TrpGlnValSerLeuAlaGpHeLysLeuAsnYtYrTrpIleHisPheCysGlyGlySerLeu	Query Match: 74.90%					
118 ATGCCACCCCGAGTGGTCTCAGCGCGCGCGCGTGGCGAGCCGAGTCAAGGATCTG						

```

Dh 65 ILeHisProGIntPrValLeuThraAlaHisCysValGlyProHisILeLysSerPro 84
Qy 178 GCCGGCCCTGAGGTCAGACCTGCGGAGCAGCACCCCTACTACTACAGACAGCTGCTGCCG 237
Dh 85 GlnLeuPheArgValGlnLeuArgGlnIntLeuTYTYTYGlyAspGlnLeuLeuSer 104
Qy 238 GTCCAGCAGATCATCTGTCACCCACAGTTCATACACCGCCGACATGCGAGCAGATCCGCC 297
Dh 105 LeuAsnArgILeValHisProHisTYrTYrThraLagLugLylValAspAla 124
Qy 298 CTGCTGGAGCTGGAGGAGCCGGTGAAGTCTTCACGCCACGTCACAGCGTCACTGCC 357
Dh 125 LeuLeuGlnLeuGlnValAlaProValAsnValSerThrHisIleHisProILeSerLeuPro 144
Qy 358 CCTGCTCAGAGACCTTCCTCCCGCGGAGTGCCTGCTGGTCACAGCTGGGCGCATGTG 417
Dh 145 ProAlaSerGlnThrPheProProGlyThrSerCysTrpValThrGlyTrpGlyAspIle 164
Qy 418 GACATATGAGAGCCCTCCACCGCCATTCCTCTGAAAGCAGAGTGAAGTCCCATTAATG 477
Dh 165 AspAsnAspGluProLeuProProProTyrrProLeuLysGlnValValProIleVal 184
Qy 478 GAAACCCACATTTGTGACGCAAAATACACCTTGCGCCCTACAGCGGAGAGAGCTCCGC 537
Dh 185 GlnAsnSerLeuCysAspArgLysThrHisThrGlyLeuTYrThrGlyAspAspPhePro 204
Qy 538 ATCGCTCCGTCAGACATCTGTCGCGGAGAACCCGAGGAGCATCTGCACGCGGCGAC 597
Dh 205 ILleValHisAspGlyMetLeuCysAlaGlyAsnThrArgAspSerCysGlnIleLysp 224
Qy 598 TCCGAGAGGCCCCCTGCTGTCGAAGGTGAATGGACACTGGCTGCACAGCGCGGTGTCAGC 657
Dh 225 SerGlyGlyProLeuValCysLysValLysGlyThrTrpLeuGlnAlaGlyValAlaSer 244
Qy 658 TGGGCGGAGGCGCTGTGCGCCACCGCCAGCGGCTGGCATCTACACCGTGTCACTACTAC 717
Dh 245 TrpGlyGlnGlyCysAlaGlnProAsnLysProGlyIleTYrThrArgValThrTYrTYr 264
Qy 718 TTGAGCTGATTCACCACTATGATGCCCAAA 747
Dh 265 LeuAspTrpILeHisArgTYrValProGlu 274

RESULT 5
TRYT_CANFA
ID TRYT_CANFA STANDARD; PRT: 275 AA.
AC PI5944:
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DB Trypsinase precursor (EC 3.4.21.59).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89352460; PubMed=2504277;
RA Vanderslice P., Craik C.S., Nadel J.A., Caughey G.H.;
RT "Molecular cloning of dog mast cell tryptase and a related protease:
RT structural evidence of a unique mode of serine protease activation.";
RL Biochemistry 28:4148-4155(1989).
CC -1- FUNCTION: TRYPTASE IS THE MAJOR NEUTRAL PROTEASE PRESENT IN MAST
CC CELLS AND IS SECRETED UPON THE COUPLED ACTIVATION-DEGRANULATION
CC RESPONSE OF THIS CELL TYPE.
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC with more restricted specificity than trypsin.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: RELEASED FROM THE SECRETORY GRANULES UPON
CC MAST CELL ACTIVATION.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. TRYPTASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL: M24664; AAA30854.1; -.
DR PIR: A32410; A32410.
DR HSSP: P20231; IMAO.
DR MEROPS: S01.143; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser_Protease_Try.
DR Pfam: PF00089; trypsin.1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRYP_SPE; 1.
DR PROSITE: PS50240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; 1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
KW Hydrolyase; Serine protease; Signal; Glycoprotein; Zymogen.
FT SIGNAL 1 20
FT PROPEP 21 30 ACTIVATION PEPTIDE (BY SIMILARITY).
FT CHAIN 31 275 TRYPTASE.
FT ACT_SITE 74 74 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 224 224 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 59 75 BY SIMILARITY.
FT DISULFID 155 230 BY SIMILARITY.
FT DISULFID 188 211 BY SIMILARITY.
FT DISULFID 220 248 BY SIMILARITY.
FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 275 AA; 30088 MW; C3B869251F248D5B CRC64;

Alignment Scores:
Pred. No.: 6,13e-56 Length: 275
Score: 1077.00 Matches: 189
Percent Similarity: 85.71% Conservative: 21
Best local Similarity: 77.14% Mismatches: 35
Query Match: 73.87% Indels: 0
DB: 1 Gaps: 0

US-09-598-982-20 (1-771) x TRYT_CANFA (1-275)
Qy 19 ATCGTGGGGGTGAGAGCCGCCAGGACAAAGTGGCCCTGAGAGTGAAGAGTCTG 78
Dh 31 ILleValGlnGlyArgValGlnAlaProGlySerLysTrpProTrpGlnValAlaSerLeuArgLeu 50
Qy 79 CAGGCGCCATCTGATGATGACTTCTGCGGGGCTCCCTCATTCACACCCCAATGGTGTG 138
Dh 51 LysGlnGlnTYrTYrPArgHisILeCysGlyLysSerLeuILeHisProGIntPrValLeu 70
Qy 139 ACCGCGCGGCGGTGGTGGAGCCGAGCATCAAGATCTGCGCGCCGCTCAGGAGTCACTG 198
Dh 71 ThrAlaAlaHisCysValGlyProAsnValAlaLysProGlnGlnIleArgValGlnLeu 90
Qy 199 CGGAGCAGCACCCTCTACTACAGAGCAGCTGCTCCGCTCAGCAGAGATCATCTGTCAC 258
Dh 91 ArgGlnGlnHisLeuTYrTYrGlnAspHisLeuLeuProValAsnArgILeValMetHis 110
Qy 259 CCACAGTTCTACACCGCCAGATGAGAGCGACATCGCCGCTGTGAGAGTGGAGAGCCG 318
Dh 111 ProAsnTYrTYrThrProGlnAsnGlyAlaAspIleAlaLeuLeuGlnLeuAspPro 130
Qy 319 GTGAAGTCTCCACCGACCGTCACAGGTCACCGTCCGCTGCTCAGAGACCTGCC 378
Dh 131 ValAsnValSerAlaHisValGlnProValThrLeuProProAlaLeuGlnThrPhePro 150
Qy 379 CCGGGATGATCCGCTGCTGCTACTGCTGCTGGCGCATGTGACATGATGAGACCGCTCCA 438
Dh 151 ThrGlyThrProGlyStrpValThrGlyTrpGlyAspValHisSerGlyThrProLeuPro 170
Qy 439 CCGCATTTCTCTGAGACAGGTAAGTCCCATTAATGAAGAAACACCATTTGTGAGCCA 498
Dh 171 ProProPheProLeuLysGlnValLysValProIleValGlnAsnSerMetCysAspVal 190

```


ID MCT7_MOUSE STANDARD: PRT: 273 AA.
 AC 002844;
 DT 01-JUN-1993 (Rel. 26, Created)
 DT 01-JUN-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Mast cell protease 7 precursor (EC 3.4.21.59) (MMP-7) (Trypsinase).
 GN MCTP7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=93087489; PubMed=1454796;
 RC MEDLINE=93087489; PubMed=1454796;
 RA McNeil H.P., Reynolds D.S., Schiller V., Ghildyal N., Gurley D.S.,
 RA Austen K.F., Stevens R.L.;
 RT "Isolation, characterization, and transcription of the gene encoding
 RT mouse mast cell protease 7.";
 RT Proc. Natl. Acad. Sci. U.S.A. 89:11174-11178(1992).
 RL [2]
 RN SEQUENCE FROM N.A. (TRUNCATED ISOFORM).
 RC STRAIN=C57BL/6J;
 RC MEDLINE=96162035; PubMed=8576265;
 RA Hunt J.E., Stevens R.L., Austen K.F., Zhang J., Xia Z.,
 RA Ghildyal N.;
 RT "Natural disruption of the mouse mast cell protease 7 gene in the
 RT C57BL/6 mouse.";
 RT J. Biol. Chem. 271:2851-2855(1996).
 RL -1- FUNCTION: TRYPSIN IS THE MAJOR NEUTRAL PROTEASE PRESENT IN MAST
 CC CELLS, AND IS SECRETED UPON THE COUPLED ACTIVATION-DEGRANULATION
 CC RESPONSE OF THIS CELL TYPE.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
 CC with more restricted specificity than trypsin.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
 CC truncated form; are produced by alternative splicing. The
 CC alternative splicing event is due to a G to A point mutation at
 CC the exon 2/intron 2 splice site and causes loss of protein
 CC expression. The alternatively spliced transcript is only found in
 CC C57BL/6 mouse.
 CC -1- DEVELOPMENTAL STAGE: IS NOT EXPRESSED IN MATURE SEROAL OR
 CC MUCOSAL MAST CELLS AND IS EXPRESSED ONLY TRANSIENTLY AT AN EARLY
 CC STAGE OF IN VITRO MAST CELL DIFFERENTIATION.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. TRYPSIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L00654; AAA39993.1; -
 DR EMBL: L00653; AAA39992.1; -
 DR EMBL: U42405; AAA97874.1; -
 DR EMBL: U42406; AAA97875.1; -
 DR PIR: A47246; A47246.
 DR HSSP: P20231; IAAO.
 DR MEROPS: S01.026; -
 DR MGD: MGI:96943; MCTP7.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR001254; Ser.protease_Try.
 DR Pfam: PF00089; trypsin.1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR SMART: SM00020; TRYPSIN_Spec.1.
 DR PROSITE: PS50240; TRYPSIN_DOM.1.
 DR PROSITE: PS00134; TRYPSIN_HIS.1.
 DR PROSITE: PS00135; TRYPSIN_SER.1.
 KM Hydrolase; Serine protease; Signal; Zymogen; Alternative splicing;
 KM Glycoprotein.
 FT SIGNAL 1 18 POTENTIAL.
 FT PROPEP 19 28 ACTIVATION PEPTIDE (POTENTIAL).

FT CHAIN 29 273 MAST CELL PROTEASE 7.
 FT ACT_SITE 72 72 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 119 119 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 222 222 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 57 73 BY SIMILARITY.
 FT DISULFID 153 228 BY SIMILARITY.
 FT DISULFID 186 209 BY SIMILARITY.
 FT DISULFID 218 246 BY SIMILARITY.
 FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 44 46 VSL -> GCC (IN TRUNCATED ISOFORM).
 FT VARSPLIC 47 273 MISSING (IN TRUNCATED ISOFORM).
 FT VARSPLIC 47 273
 SQ SEQUENCE 273 AA; 30337 MW; 50ECB4765294205E CRC64;

 Alignment Scores:
 Pred. No.: 8.95e-55 Length: 273
 Score: 1057.00 Matches: 189
 Percent Similarity: 82.07% Conservative: 17
 Best Local Similarity: 75.30% Mismatches: 43
 Query Match: 72.50% Indels: 2
 Gaps: 1

 US-09-598-982-20 (1-771) x MCT7_MOUSE (1-273)
 QY 1 GGGCCCTCGAGAAAGA-----ATCGTGGGGGTCCAGAGGCCCGGAGGCAATG 54
 DB 21 GlyProAlaMetThrArgGluGlyLeValGlyGlnGlnAlaHisGlyAsnLysTrp 40
 QY 55 CCGTGGCAGGTGAGTCCAGAGTCCAGCGCCCATCTGATGCACTTGTGGGGGTCC 114
 DB 41 ProTrpGlnValSerLeuAlaGlnAlaHisGlyValGlyProAspValAlaAsp 60
 QY 115 CTCATCCACCCCACTGGGTGCTGACCCCGCGGCGTGGTGGACCGGAGTCAAGAT 174
 DB 61 LeuIleHisProGlnTrpValLeuThrAlaHisGlyValGlyProAspValAlaAsp 80
 QY 175 CTGGCGCGCTCAGGGGTCCAGTGGGGGAGCAGCAGCTCTACAGGACAGCGTGTG 234
 DB 81 ProAsnLysValArgValGlnLeuAlaGlnLysGlyLeuTrpValThrHisAspHisLeuMet 100
 QY 235 CCGGTACACGAGTATGTCGTCGACCCACAGTCTACACCGCCGAGTCCGAGCATC 294
 DB 101 ThrValSerGlnIleIleThrHisProAspPheTrpIleValGlnAspGlyAlaAspIle 120
 QY 295 GCGTGTGGAGTCCAGAGCCGCGTGAAGTCCACGACGTCACAGGTCACCGTG 354
 DB 121 AlaLeuLeuLysLeuThrAsnProValHisLeuSerAspTrpValHisProValProLeu 140
 QY 355 CCGCTCGCTCCAGAGACCCCTCCCGGCGGATCCGCTGCGTACAGCGGGCGCAT 414
 DB 141 ProProAlaSerGlnThrPheProSerGlyThrLeuGlyStrpValThrGlyTrpLysn 160
 QY 415 GTGACATGATGATGACCGCTCCACCGCATCTCTCTGAAACGAGTGAAGTCCCATTA 474
 DB 161 IleAspAsnGlyValAlaHisLeuProProProPheProLeuLysGlnValGlnAlaProIle 180
 QY 475 ATGGAACACCATTTGTGACGCAAAATPACACCTTGGCGGCTACACGGAGACAGCTG 534
 DB 181 IleGlnAsnHisLeuGlySerLeuLysTrpHisLysGlyLeuIleThrGlyAspAsnVal 200
 QY 535 CCGATCGTCCCGTCCAGAGTGTGTGTGCGGGAGACACCGGAGGAGCATGACCGAGGC 594
 DB 201 HisIleValAlaArgAspAspMetLeuGlySalGlyAsnGlnGlyHisAspSerGlyGlnGly 220
 QY 595 GACTCGGAGGCGCGTGGTGTGCAAGTGAATGTCACCTGCTCCAGGCGGCGTGTGTC 654
 DB 221 AspSerGlyGlyProLeuValLysLysValGlnAspThrTrpLeuGlnAlaGlyValVal 240
 QY 655 AACTGGGCGAGGCGTGTGCCAGGCCAAACGCGCTGAGCATTAACCCGTGTACCTAC 714
 DB 241 SerTrpGlyGlyGlyLysAlaGlnProAsnArgProGlyIleGlyThrArgValThrTrp 260
 QY 715 TACTTGACTGGATCCACCATATGTCCCAA 747


```

Db      261  TYLEASPTRPDLEHISHTYValProlys 271
RESULT 8
MCT6_RAT
ID      MCT6_RAT      STANDARD:      PRT:      274 AA.
AC      P50343; P97593;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Mast cell protease 6 precursor (EC 3.4.21.55) (MCP-6) (Trypsinase).
GN      MCP6 OR MCP6.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
[1]
RX      NCBI_TaxID=10116;
RP      SEQUENCE FROM N.A.
RC      TISSUE=Peritoneal mast cells;
RX      MEDLINE=96015171; PubMed=8537314;
RA      Ide H., Itoh H., Tomita M., Murakumo Y., Kobayashi T.,
RA      Maruyama H., Osada Y., Nawa Y.;
RT      "cDNA sequencing and expression of rat mast cell tryptase.";
RL      J. Biochem. 118:210-215(1995).
[2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Sprague-Dawley; TISSUE=Peritoneal mast cells;
RX      MEDLINE=97149430; PubMed=8996238;
RA      Lutzeltschew C., Fejler G., Aveskog M., Hellman L.;
RT      "Secretory granule proteases in rat mast cells. Cloning of 10
RT      different serine proteases and a carboxypeptidase A from various rat
RT      mast cell populations.";
RL      J. Exp. Med. 185:13-29(1997).
CC      -I- FUNCTION: TRYPTASE IS THE MAJOR NEUTRAL PROTEASE PRESENT IN MAST
CC      CELLS. AND IS SECRETED UPON THE COUPLED ACTIVATION-DEGRANULATION
CC      RESPONSE OF THIS CELL TYPE.
CC      -I- CATALYTIC ACTIVITY: Preferential cleavage: Arg-I-, Lys-I-, but
CC      with more restricted specificity than trypsin.
CC      -I- SIMILARITY: BELONGS TO TRYPTASE FAMILY SL. TRYPTASE SUBFAMILY.
CC
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
CC
DR      EMBL: D38455; BAA07486.1; -.
DR      EMBL: U67909; AAB48262.1; -.
DR      HSSP: P20231; IAOO.
DR      MEROPS: S01.025; -.
DR      InterPro: IPR001314; Chymotrypsin.
DR      InterPro: IPR001254; Ser-protease_Try.
DR      Pfam: PF00089; trypsin; 1.
DR      PRINTS: PR00722; CHYMOTRYPSIN.
DR      SMART: SM00020; TRYP-SPEC. 1.
DR      PROSITE: PS50240; TRYPsin DOM. 1.
DR      PROSITE: PS00134; TRYPsin_HIS. 1.
DR      PROSITE: PS00135; TRYPsin_SER. 1.
KW      Hydrolase; Serine protease; Glycoprotein; Multigene family;
KW      Signal.
FT      SIGNAL. 1 19
FT      PROPEP 20 29 POTENTIAL.
FT      CHAIN 30 274 ACTIVATION PEPTIDE.
FT      ACT_SITE 73 73 MAST CELL PROTEASE 6.
FT      ACT_SITE 120 120 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      ACT_SITE 223 223 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      ACT_SITE 223 223 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT      DISULFID 58 74 BY SIMILARITY.
FT      DISULFID 154 229 BY SIMILARITY.
FT      DISULFID 187 210 BY SIMILARITY.
FT      DISULFID 219 247 BY SIMILARITY.
FT      CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT      CARBOHYD 131 131 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CONFLICT 128 128 I -> N (IN REF. 2).
FT      CONFLICT 139 139 I -> T (IN REF. 2).
SQ      SEQUENCE 274 AA; 30508 MW; DF84DB55668CA1A25 CRC64;

Alignment Scores:
Pred. No.: 1,17e-54 Length: 274
Score: 1055.00 Matches: 187
Percent Similarity: 84.06% Conservative: 24
Best Local Similarity: 74.50% Mismatches: 38
Query Match: 72.36% Indels: 2
DB: 1 Gaps: 1

US-09-598-982-20 (1-771) x MCT6_RAT (1-274)
QY      4 CCCCCTCGAAGAAAG-----ATCGTCGGGGGCTCAGAGAGCCSSCAGAGCAAGTGGCC 57
Db      23 ProValLysGlnArgValGlcIylLeValGlcIylArgGlnLysSerGlnSerLysTyrPro 42
QY      58 TGGCAGGTGAGCGCTGAGATGCCAGCGCCCATCTGAGTGCATCTGCGGGGGCTGGCTC 117
Db      43 TrpGlnValSerLeuArgPheLysPheSerPheTrpMetHisPheCysGlyLysSerLeu 62
QY      118 ATSCAASSCCAGTGGGTGCTGACCGCGCGGCTGCTGGAGCGGAGCTCAAGATCTG 177
Db      63 LlenIsProGlnTrpValLeuThrAlaLanHisCysValGlyLeuHisIleLysSerPro 82
QY      178 GCGCGGCTCAGGCTGACGTCGCGGAGAGCAAGCTGTACACAGAGAGAGAGAGCTGCGG 237
Db      83 GluLeuPheArgValGlnLeuArgGlnGlnIylLeuLysTyrAlaLysPheLeuLeuThr 102
QY      238 GTGACGAGCATCATGCTGTCGACCCAGCATGCTGACCGCGGCTGACGAGCGGAGCATGCGC 297
Db      103 ValAsnArgThrValValHisProHisTyrTrpThrValGlnAspGlyValAspPheLeu 122
QY      298 TGTGTGAGCTGAGAGAGAGCGGCTGAGAGCTGTCAGACAGCTGACAGAGCTGCGCC 357
Db      123 LeuLeuGlnLeuGlnIleProValAsnValSerThrHisIleHisProLysSerLeuPro 142
QY      358 CCTGCTCAGAAACCTTCCCGCGGGGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 417
Db      143 ProLysSerGlnThrPheProSerGlnThrSerCysTyrValThrGlyTyrPheLysPhe 162
QY      418 GACATGATGAGAGCGGCTGCCAGCGCATCTGCTGAGAGAGAGTGAAGTCCCGCATATG 477
Db      163 AspSerAspGlnProLeuLeuProTyrProLeuLysGlnValLysValProLysVal 182
QY      478 GAAACACATTTGTGACGCAAAATACACACTTGGCGGCTACACGAGAGAGAGAGCTCGGC 537
Db      183 GluAsnSerLeuCysAspArgLysTyrHisThrGlyLeuTyrThrGlyLysPheValPro 202
QY      538 ATGCTGCCGTCGAGACATGCTGTGTGCTGCGGAAACACCGGAGGAGCATCATGCCAGGCGGC 597
Db      203 IleValGlnAspGlyMetLeuCysAlaGlyAsnThrArgSerAspSerCysGlnLysPhe 222
QY      598 TCCGAGAGGCGGCTGCTGTCGCAAGGAGTACGACCTGCTGCTGCTGCTGCTGCTGCTG 657
Db      223 SerGlyGlyProLeuValCysLysValLysGlyThrTrpLeuGlnAlaLysValSer 242
QY      658 TGGGAGAGAGGCTGCTGCCAGCGCCAGCGGCTGAGCATCTACACCGGCTGCTGCTGCTGCTG 717
Db      243 TrpGlyGlnGlyCysAlaGlnLysAlaAsnArgProGlyIleTyrThrArgValThrGlyTyr 262
QY      718 TTGAGACTGATTCACACACATATGCCCCAAAAG 750
Db      263 LeuAspTrpLlenHisArgTyrValProGlnArg 273

RESULT 9
MCT7_RAT
ID      MCT7_RAT      STANDARD:      PRT:      273 AA.
AC      P27435; P27436;
DT      01-AUG-1992 (Rel. 23, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)

```


RA Yano M., Yang B., Kido H.:
RT "Mast cell tryptase from pig lungs triggers infection by pneumotropic
Sendai and Influenza A viruses. Purification and characterization.";
RL Eur. J. Biochem. 267:3189-3197(2000).
CC
CC -1 FUNCTION: TRYPTASE IS THE MAJOR NEUTRAL PROTEASE PRESENT IN MAST
CC CELLS AND IS SECRETED UPON THE COUPLED ACTIVATION-DEGRADULATION
CC RESPONSE OF THIS CELL TYPE.
CC
CC -1 CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC with more restricted specificity than trypsin.
CC
CC -1 SUBUNIT: HOMOTETRAMER (By similarity).
CC
CC -1 SUBCELLULAR LOCATION: RELEASED FROM THE SECRETORY GRANULES UPON
CC MAST CELL ACTIVATION.
CC
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1, TRYPTASE SUBFAMILY.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

	CC	
DR	-:::	EMBL; AB038652; BAA93614.1; -.
DR	HSSP;	P20231. IAOO.
DR	MEROPS;	S01_143;-.
DR	InterPro;	IPIR001314; Chymotrypsin.
DR	InterPro;	IPIR001254; Ser_protease_Try.
DR	Pfam;	PF00089; trypsin; 1.
DR	PRINTS;	PRO0722; CHYMOTRYPSIN.
DR	SMART;	SM00020; Tryp_SPC; 1.
DR	PROSITE;	PS50240; TRYPSIN_DOM; 1.
DR	PROSITE;	PS00134; TRYPSIN_HIS; 1.
DR	PROSITE;	PS00135; TRYPSIN_SER; 1.
KW	Hydrolase;	Serine protease; signal; glycoprotein; zymogen.
FT	SIGNAL	1 20
FT	PROPEP	CHAIN
FT	CHAIN	21 30
FT	ACT_SITE	74 74
FT	ACT_SITE	121 121
FT	ACT_SITE	224 224
FT	DISEULFID	59 75
FT	DISEULFID	155 230
FT	DISEULFID	188 211
FT	DISEULFID	220 248
FT	CARBOHYD	132 132
FT	CARBOHYD	233 233
QO	SEQUENCE	275 AA; 30439 MW; ACC582647EFCB973 CRC64; N-LINKED (GLCNAC...) (POTENTIAL). N-LINKED (GLCNAC...) (POTENTIAL).

Alignment Scores:	
Pred. No.:	8.31e-52
Score:	1005.00
Percent Similarity:	82.38%
Best Local Similarity:	74.18%
Query Match:	69.00%
OB:	1
	0
	Gaps: 0
	Indels: 0
	Mismatches: 43
	Conservative: 20
	Matches: 181
	Length: 275

US-09-598-982-20 (1-771) x TRYF_PIG (1-275)

[illegible]

OY	259	CCACAGTTCTACACCGCCCAATGGAGCCGACATGCCCTGCTGGACGCTGGAGAGACCG	318
Db	111	ProsnryttrypasgIvalaIasnqIYalaSpllealeuIeugIuIeugIuaspIe	130
OY	319	GTAAGGCTCTCCACCCACGTCACACGCGTCAACCTGCCCTGCCTCAGACGCTTCC	378
Db	131	ValasnIeuserSerHIsvalaIgnIProvalIthIreupProIaIasergIuIthIrepro	150
OY	379	CCGGGAGATGCCGTCTGGGTCACTGGCTGGGGCGATGGACATATATAGACGCCCTCCA	438
Db	151	LyseIyIthIarGystrypvalIthGlyItrpelysIpaIhIsseIrglyItrpIeupIro	170
OY	439	CCGCAATTCCTCTGAAGCAGAGTGAAGCTGCCCATATATGGAAAAACCATTTCTGACGCA	498
Db	171	ProIroIytrpIeulayIsgIvalaIargValIproIlevalaIguIasnserGIucyIasImet	190
OY	499	AAATPACCACTTGGCGCCCTACACGGGAGAGAGAGCTCCGCACTGCTCGCTACAGACATGCTG	558
Db	191	GInIytrHIsIeugIyIleuSerTthGlyIasPasnIleIroIlevalaIargaspasImetIleu	210
OY	559	TGTGCGGGAGACACCCGGAGGAGACTCATGTGCCAGGGCCACATCCCGGAGGGCCCTGGTGTGC	618
Db	211	CysIaIagIyserGIuGIuIhIsAsIserCysIeInIyAsIserGIuGIyIroIeulayICys	230
OY	619	AAGTGAATGGCACCTGGCTGCTCAAGCGGGCGTGGTCAAGCTGGGGCGAGGGCTGTGCCAG	678
Db	231	ArgValaInclYrIthItrpIeugIaIaIagIyValaIasIerItrpGIuGIuGIyCysAlaIeu	250
OY	679	CCCAACCGGCTGGCAGTATACACCGGTCATCCACACTACTGTGGAGTGGATGCCACACTAT	738
Db	251	ProIasnIargIroGIyIleIytrHtrArgValaIthIhIsIyIreusIerItrIeInIsIeInCys	270
OY	739	GTCGCCCAAAAG 750	
Db	271	IleProIarGIu 274	

RESULT	12
TRYD_HUMAN	
ID_TRYD_HUMAN	STANDARD;
AC	PRT; 235 AA.
Q95BZ3; Q9H2Y6; Q95824;	
16-OCT-2001 (Rel. 40, Created)	
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Putative trypsinase delta precursor (EC 3.4.21.-) (Mast cell mMCP-7-like) (mMCP-3-like trypsinase III).
DE	TPSD1.
GN	

OS Homo sapiens (Human).
OC Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP
RP SEQUENCE FROM N.A.
RX MEDLINE=99121069; Pubmed=9920877;
RA Pallao M., Fejzo M.S., Shaysteh L., Blount J.L., Caughey G.H.;
RT "Characterization of genes encoding known and novel human mast cell
RT tryptases on chromosome 16p13.3.";
RL J. Biol. Chem. 274:3355-3362(1999).
RN [2]
RP
RP SEQUENCE FROM N.A.
RX MEDLINE=2110154; Pubmed=11174199;
RA Min H.K., Kambe N., Schwartz L.B.;
RT "Human mouse mast cell protease 7-like tryptase genes are
RT pseudogenes.";
RL J. Allergy Clin. Immunol. 107:315-321(2001).
RN [3]
RP
RP SEQUENCE OF 46-235 FROM N.A.
RX Wang H.W., McNeill H.P., Thomas P.S., Murphy B.N., Webster M.J.,
RA Hettiaratchi A., King G., Heywood G.J., Huang C., Stevens R.L.,
RA Hunt J.E.;
RT "Molecular cloning and characterization of novel human tryptase cDNAs
RT and splicing variants.";
RL submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.

Alignment Scores:					
Pred. No. :	9	37e-51	Length:	235	
Score:	988.00		Matches:	173	
Percent Similarity:	91.71%		Conservative:	15	
Best Local Similarity:	84.39%		Mismatches:	17	
Query Match:	67.76%		Indels:	0	
DB:	1		Gaps:	0	
US-09-598-982-20 (1-771) x TRYD_HUMAN (1-235)					
QY	19	ATCGTCGGGGSCTCAGAGAGCCCCCAGAGACGATGCCCTTGCCAGGTGCCTGAGAGTC	78		
Db	31	IleValGIgylgIngluaIProARSerlyStrPrOTrgInValserLeuArgAl	50		
QY	79	CACGGCCCATCTAGTGGATGCACTTCCTGGGGGGGCTCCGCATCACCCCCAGTGGCTGTG	138		
Db	51	ArgGIyProTyrrTrmehNHisPhcysGIyLysertLeuIleHISproGINtrPvalLeu	70		
QY	139	ACCGCCGCGSGCTCGTGGGACGCGGAGCGTCAAGGATGTGCCGCGCTCAGGGTCAAATG	198		
Db	71	ThrlAalaINAsKcySValIGluProASPIleyLysaspLeuAlalaleuArgValGIleu	90		
QY	199	CGGGAGCAGCACTCTATTACTACGAGCAGCATGCTGGCGGACACAGATCATGTGTCAC	258		
Db	91	ArgGIuGlInHISLeuTyrrTYrGIlnasPGLnLeuLeuProVALserArgIIleIleValHIS	110		
QY	259	CCACAGTTCTACACCGCCAGATCGGAGCGGACATCGCCCTGCTGAGCTGGAGAGCCG	318		
Db	111	ProGIInPheTyrrIleIleGIInThGIYAlasPlleAlaleuNGluLeuGIluPro	138		
QY	319	GTGAAGGCTCTCAGCAGAGCTCAGACAGGTATCACTGGCCCCCTGCTCAGAGACTTGCCC	378		
Db	131	ValasnIleSerSerHISIHISIThrValThrLeuProProlAlaserGIuThrPhePro	150		

[illegible]

```

FT CARBOHYD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 269 AA; 29824 MW; BAA4AF7E4E2FD4E5 CRC64;

Alignment Scores:
Pred. No.: 4.29e-33 Length: 269
Score: 684.50 Matches: 132
Percent Similarity: 65.34% Conservative: 32
Best Local Similarity: 52.59% Mismatches: 78
Query Match: 46.95% Indels: 9
DB: 1 Gaps: 4

US-09-598-982-20 (1-771) x TRYX_MOUSE (1-269)

OY 19 ATCTGCGGGGTCAGAGAGCCCGGAGAGCAAGTGGCCCTGAGCTGAGAGTTC 78
D 19 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 20 lIeValGlGyGyCyAlValProAlaArgYrProTrpGlnValSerLeuArgPhe 39
OY 79 CAGGCG-----CCAACTGATGCACTTTCGCGGGGCTCCCTCATCCACCCCGAG 129
D 79 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 40 HIsGlMeTGlYserGlYglnTrpGlnHisIleCysGlYglnSerLeuIleHisProGln 59
OY 130 TGGGTCGAGACCGCGCGCGCGCGCGGAGACCGGACGTCAGATCTGGCGCGCGTCAAG 189
D 130 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 60 TrpValIleuThAlaAlaHisCysValGluleuGluGlYleuGluAlaAlaThrIleuArg 79
OY 190 GTGCACTGCGGCGGAGACACCTTACTACTACAGAGACAGTCTGCTCGGTCAGACAGATC 249
D 190 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 80 ValGlnValGlYglnLeuArgIleuArgYrAspHisAspGlnLeuCysAsnValThrGluIle 99
OY 250 ATCTGTCACCCACAGTCTTACACCGCCGAGATGGA-----GCGGACATCGCCCTTC 300
D 250 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 100 lIeArgHisProAsnPhaAsnMetSerTrpYrGlYTrpAspThrAlaAspIleAlaLeu 119
OY 301 CTGGAGCGGAGAGACCGGTCGAGAGTCTCCACAGCGTCCACAGCTGACCGTCCCGCCCT 360
D 301 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 120 LeuYlStleuGluAlaProIleuThrIleuSerGluAspValAsnLeuValSerLeuProSer 139
OY 361 GCGTCAGAGACCTTCCCGCGGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
D 361 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 140 ProSerIleuIleValProProGlnIleuSerTrpValThrGlYTrpIleYlspIleAla 159
OY 421 AATGATGAGCGGCTCCCGACCGCATTTCTCTGAAAGCAGAGTCAAGGTCGCCATATGGA 480
D 421 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 160 AsnHisThrProIleuProProProTrpYrHisIleuGlnGluValGlYValProIleValGlY 179
OY 481 AACCAACATTTGAGCGCAAAATACCACTTGGCGGCTTACAGCGGAGAGACAGCGCCGATC 540
D 481 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 180 AsnAlaGluCysAsnGlyHisTrpGln-----ThrIleLeuGluGlnAspAspGluVal 197
OY 541 GTCCGTGAGACATGCTGTGTGCTCCGAGAACACCCGAGGAGACTCATCCAGGCGGACATCC 600
D 541 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 198 lIeYlStGlnAspMetIleuSerGlnIleuSerGlnIleuSerGlnIleuSerGlnIleuSer 217
OY 601 GAGAGGCGGCTGCTGTCAGAGTGAATGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 660
D 601 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 218 GlYglYrProIleuValCysArgTrpYrCysTrpIleGlnValGlYValIleSerTrp 237
OY 661 GCGGAGGCGGCTGCGGAGCGGCTGAGCTGATACCGCGGCTGAGCTGAGCTGAGCTGAG 720
D 661 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 238 GlYTrpGlyGyGlyTrp---AsnLeuProGlnValTrpAlaArgValThrSerTrpVal 256
OY 721 GACTGATTCACCACTATGATCCCAAAAAGCCG 753
D 721 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 257 SerTrpIleHisGlnHisIleProIleuSerPro 267

RESULT 14
TRYX_MOUSE STANDARD: PRT: 311 AA.
AC 090UL7:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE Trypsin gamma precursor (EC 3.4.21.-) (Transmembrane trypsin).
GN TP5G1 OR TMT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv, and BALB/C;
RC MEDLINE=99452974; PubMed=10521469;
RA Wong G.W., Tang Y., Feyfant E., Sali A., Li L., Li Y., Huang C.,
RA Friend D.S., Krilis S.A., Stevens R.L.;
RT "Identification of a new member of the trypsin family of mouse and
RT human mast cell proteases which possesses a novel COOH-terminal
RT hydrophobic extension."
J. Biol. Chem. 274:30784-30793 (1999).
CC -!- SUBCELLULAR LOCATION: MEMBRANE-ANCHORED (Potential).
CC -!- TISSUE SPECIFICITY: Expressed in many tissues.
CC -!- SIMILARITY: BELONGS TO TRYPSIN SUBFAMILY.
CC -!- SIMILARITY: BELONGS TO TRYPSIN SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF175760; AAF03698.1; -
DR EMBL: AF175523; AAF03696.1; -
DR HSBP: P20231; IAAO.
DR MEROPS: S01.028; -.
DR MGD: MGI:1349391; Tpsgl.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser-protease_Try.
DR Pfam: PF00089; trypsin.1
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRYP-SPC.1.
DR PROSITE: PS00240; TRYPSIN_DOM.1.
DR PROSITE: PS00134; TRYPSIN_HIS.1.
DR PROSITE: PS00135; TRYPSIN_SER.1.
DR K1 Hydroxylase: Serine protease; Signal; Glycoprotein; Zymogen;
KM Transmembrane.1
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 28 TRYPSIN GAMMA LIGHT CHAIN.
FT CHAIN 30 311 TRYPSIN GAMMA HEAVY CHAIN.
FT TRANSMEM 277 297 POTENTIAL.
FT ACT_SITE 70 70 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 117 117 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 214 214 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 18 137 INTERCHAIN (POTENTIAL).
FT DISULFID 55 71 BY SIMILARITY.
FT DISULFID 151 220 BY SIMILARITY.
FT DISULFID 184 202 BY SIMILARITY.
FT DISULFID 210 238 BY SIMILARITY.
FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 32656 MW; 7FC9D6F6A2A808 CRC64;
SQ SEQUENCE 311 AA; 32656 MW;

Alignment Scores:
Pred. No.: 1.13e-28 Length: 311
Score: 608.50 Matches: 123
Percent Similarity: 60.63% Conservative: 31
Best Local Similarity: 48.43% Mismatches: 81
Query Match: 41.74% Indels: 19
DB: 1 Gaps: 6

US-09-598-982-20 (1-771) x TRYX_MOUSE (1-311)

OY 16 AGAATGTCGCGGTCAGAGAGCCCGGAGAGCAAGTGGCCCTGAGCTGAGAGTTC 75
D 16 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
D 29 ArgIleValGlYglnHisAlaProAlaGlyThrTrpProTrpGlnIleAspSerLeuArg 48

```

```

OY 76 CTCACGGCCCATCTACTGATGCTTTCGCGGGGCTCCCTCATCCACCCCGCATGGGTG 135
    ::::: ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 49 LeuH1sLys-----ValHisValCysGlyGlySerLeuLeuSerProLutrrPVal 65
OY 136 CTGACCCCGCGCGCTGGTGGGACCGGACGTCAAGATCTGCGCGCTCATGGGTGCAA 195
    ||||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66 LeuH1rAlaAlaHisCysPheSerGlySerValAsnSer---SerAspTrcInValHis 84
OY 196 CCGCGGGGACACCATCTACTACAGAGACAGCTGCTCCGCTGACGAGCTGTCATCGG 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85 LeuGlyGlyLeuThrValThrLeuSerProHisPheSerThrValLysTrgIleIleMet 104
OY 256 CACCCACAGTTCTACACCGCCCGATCGGA-----GCGGACATCCGCGCTG 300
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 105 -----TyrThrGlySerProGlyProGlySerSerGlyLysPleAlaLeu 120
OY 301 CTGAGAGCTGGAGAGCGGTGAAGTCTTCACGACGTCACACGGTCAACCTGCCCT 360
    ::::: ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 ValGlnLeuSerSerProValAlaLeuSerSerGlnValGlnProValCysLeuProGln 140
OY 361 GCGTCAAGACCTTCCCGCGGGGATGCCGTGCTGGTCACTGGCGGGCGCATGTGAC 420
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141 AlaSerAlaAspPheTyrProGlyMetGlnCysTTrpValThrGlyTrpGlyThrGly 160
OY 421 AATGATGAGCGGCTCCACCGCATTTCTCTGAAGCAGGTGAAGTCCCATTAATGAA 480
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 161 GlnGlyGlnProLeuLysProProTrpAsnLeuGlnAlaLysValSerValValAsp 180
OY 481 AACCAATTGTGAGCGCAAAATACCACTTGGCGCTACACGGGAGACGACGTCCGCATC 540
    ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 ValLysThrCysSerGln-----AlaTrpAsnSerProAsnGlySerLeu 195
OY 541 GCCCGGAGACGCTGCTGTCGCGGGAGACACCGGAGGAGCTGACGAGCGCGCATCC 600
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 196 IlegInProAspMetLeuGlySalaArgGlyPro---GlyAspAlaCysGlnAspSer 214
OY 601 GGAGGGCCCTGCTGTGCAAGGTGAATGGACCTGCTGACGCGCGCGCTGTGAGCTGG 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 215 GlyGlyProLeuValCysGlnValAlaGlyThrTrpGlnAlaGlyValValSerTrp 234
OY 661 GCGGAGGGCTGTGCCACGCCGCGCTGCACTTACACCCGCTGCTCACTACTG 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 235 GlyGlyGlnCysGlyArgProAspArgProGlyValTyrAlaArgValThrAlaTyrVal 254
OY 721 GACTGGATCCACCATCTCCCAAAAGCCGTGAAGGGC 762
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 255 AsnTrpIleHisHisIleProGlnAlaGlyGlySerGly 268

RESULT 15
ID TRYG_HUMAN STANDARD: PRT; 321 AA.
AC O9NR02: O9NR08: O9C015: Q9UBB2:
DT 16-OCT-2001 (rel. 40, created)
DT 16-OCT-2001 (rel. 40, last sequence update)
DT 16-OCT-2001 (rel. 40, last annotation update)
DE Trypase gamma precursor (EC 3.4.21.-) (Transmembrane tryptase).
GN TP5G1 OR TMT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (VARIANTS GAMMA-1 AND GAMMA-2).
RX MEDLINE=20302813; PubMed=10843716;
RA Caughey G.H., Raymond W.W., Blount J.L., Hau L.W., Pallaro M.,
RA Wolters P.J., Verghese G.M.;
RT "Characterization of human gamma-tryptases, novel members of the
RT chromosome 16p mast cell tryptase and prostasin gene families.";
RL J. Immunol. 164:6566-6575(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=9452974; PubMed=10521469;
RA Wong G.W., Tang Y., Feyfant E., Sali A., Li L., Li Y., Huang C.,

```

```

RA Friend D.S., Krilis S.A., Stevens R.L.;
RT "Identification of a new member of the tryptase family of mouse and
RT human mast cell proteases which possesses a novel COOH-terminal
RT hydrophobic extension.";
RN J. Biol. Chem. 274:30784-30793(1999).
RL [3]
RP SEQUENCE OF 220-321 FROM N.A.
RA Mitten S., Agnew W.S.;
RT "Organization and alternative splicing of CACNA1H.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -I- SUBCELLULAR LOCATION: MEMBRANE-ANCHORED (Potential).
CC -I- TISSUE SPECIFICITY: Expressed in many tissues.
CC -I- POLYMORPHISM: There are two alleles; gamma-I and gamma-II which
CC differ by 5 residues.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. TRYPTASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, AF191031; AAF76457.1; -
DR EMBL, AF195508; AAF76458.1; -
DR EMBL, AF175759; AAF03697.1; -
DR EMBL, AF175522; AAF03695.1; -
DR EMBL, AF223563; AAG48852.2; -
DR HSPD, P00763; IDPO.
DR MEROPS, S01.028; -.
DR GeneID, HGNC:14134; TP5G1.
DR InterPro, IPR001314; Chymotrypsin.
DR InterPro, IPR001254; Ser-protease_Try.
DR Pfam, PF00089; trypsin; 1.
DR PRINTS, PR00722; CHYMOTRYPSIN.
DR SMART, SMO0020; TRYP_SPE; 1.
DR PROSITE, PS500240; TRYPSIN_DOM; 1.
DR PROSITE, PS00135; TRYPSIN_SER; FALSE_NEG.
DR PROSITE, PS00135; TRYPSIN_SER; FALSE_NEG.
DR Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen;
KW Transmembrane; Polymorphism.
FT SIGNAL 1 19
FT CHAIN 20 36
FT CHAIN 38 321
FT TRANSMEM 284 304
FT ACT_SITE 78 78
FT ACT_SITE 125 125
FT ACT_SITE 222 222
FT DISULFID 26 145
FT DISULFID 63 79
FT DISULFID 159 228
FT DISULFID 192 210
FT DISULFID 218 246
FT CARBOHYD 85 85
FT VARIANT 60 60
FT VARIANT 126 126
FT VARIANT 132 132
FT VARIANT 204 204
FT VARIANT 288 288
FT CONFLICT 160 160
FT SEQUENCE 321 AA; 33827 MW; FEF7B06E3C4A962D CRC64;

Alignment Scores:
Pred. No.: 8.39e-28
Score: 593.50
Percent Similarity: 61.89%
Best Local Similarity: 48.77%

Length: 321
Matches: 119
Conservative: 32
Mismatches: 80

```


Query Match: 40.71% Indels: 13
DB: 1 Gaps: 6

US-09-598-982-20 (1-771) x TRYG_HUMAN (1-321)

```
OY 16 AGAATCGTGGGGGTGACGAGGCCCCCAGAGCAAGTGGCCCTGGCAGGTGAGCTGAGA 75
    |||||||
Db 37 ArgileValIglYlSerAlaProAlaIglYalTrpGlnAlaSerLeuArg 56
OY 76 GTCCACGGGCCCATCTACTGATGACATTCTCGGGGGCTCCCTCATCCACCCAGTGGGTG 135
    :::
Db 57 LeuArgArg-----MetHisValCysGlyGlySerLeuLeuSerProGlnTrpVal 73
OY 136 CTGACCCGCGCGGCGTGGCTGGAGACCGGAGCTCAAGATCTGGCCCTCAGGGGTGCAA 195
    |||||||
Db 74 LeuThrAlaAlaHisCysPheSerGlySerLeuAnsSer---SerAspTyrGlnValHis 92
OY 196 CTGCGGGAGACAGCAGCTTACTACCGAGACAGCTGCTGCCGTACAGCATCATCGTG 255
    || |||
Db 93 LeuGlyGluLeuGluIleThrLeuSerProHisPheSerThrValArgGlnIleLeu 112
OY 256 CACCCACAGTTTACACCGCCAGATCGGA-----GGGACATCGCCCTGCTGGAGCTG 309
    |||
Db 113 HisSerSer---ProSerGlyInProGlyThrSerGlyAspIleAlaLeuValGluLeu 131
OY 310 GAGGAGCCGCTGAAGCTTCACAGCCAGCTCCACAGGTCACCCCTGCCCTGCAGAG 369
    |||||||
Db 132 SerValProValThrLeuSerSerArgIleLeuProValCysLeuProGlnLaserAsp 151
OY 370 ACCTTCCCCCGGGGATGCGCTGCTGGTCACTGCTGGGGCGATGTGACAAATGATGAG 429
    || |||||||
Db 152 AspPheCysProGlyIleArgCysTrpValThrGlyTrpGlyTyrThrArgGluGlyLeu 171
OY 430 CGCCCTCCACCGCGCATTTCCCTGAAGCAGTGAAGTCCCATATGGAACCAACCATTT 489
    |||||||
Db 172 ProLeuProProProIyrSerLeuArgIuValLysValSerValValAspThrGluThr 191
OY 490 TGTGACGCAAAATACACCTTGCGCTACACAGGAGACGACGTCGCGATGCCGTGAC 549
    |||
Db 192 CysArgArgAsp-----TyrProGlyProGlyGlySerIleLeuGlnPro 206
OY 550 GACATGCTGTGTGCGGGAGACACCGGAGGAGCTCATGCCAGGGGAGCTCCGAGGGCCC 609
    |||||||
Db 207 AspMetLeuCysAlaArgGlyPro--GlyAspAlaCysGlnAspAspSerGlyGlyPro 225
OY 610 CTGGTGTGCAAGTGAATGGCACTGGCTGACGAGGGGCGGTGTCAGCTGGGGCGAGGC 669
    |||||||
Db 226 LeuValCysGlnValAsnGlyAlaTrpValGlnAlaGlyIleValSerTrpGlyGluGly 245
OY 670 TGTGCCAGCCCAACCGGCTGCACTTACACCCGTCACCTACTTGAAGTGGATGATC 729
    |||
Db 246 CysGlyArgProAsnArgProGlyValTyrThrArgValProAlaTyrValAsnTrpIle 265
OY 730 CACCACTATGTC 741
    :::::
Db 266 ArgArgHisIle 269
```

Search completed: January 31, 2003, 06:54:08
Job time : 34.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OW nucleic - protein search, using frame_plus_n2p model

Run on: January 31, 2003, 06:53:20 ; Search time 16.5 Seconds
(without alignments)
2749.706 Million cell updates/sec

Title: US-09-598-982-20
Perfect score: 1458
Sequence: 1 gggccctcgagaagaat.....cgtgaagcggcgcgcgcgt 771

Scoring table:
BLOSUM62
Xgapop 10.0, Ygapext 0.5
Xgapop 6.0, Ygapext 7.0
Delop 6.0, Delext 7.0

Searched: 262574 seqs, 2942292 residues
Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=x1h
-Q=/cgn2_1/USPTO/spool/US09598982/runat_27012003_073718_19604/app_query.fasta.1.967
-DB=Issued_Patents_AA -QPM=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MILEN=0 -MAXLEN=2000000000
-USER=US09598982.qcgn.1.1.4.etrnat.27012003_073718_19604 -NCPI=6 -ICPU=3
-NO.XLPHY -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/6C.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1387	95.1	249	4	US-09-079-970A-5 Sequence 5, Appli
2	1368	93.8	245	4	US-09-079-970A-6 Sequence 6, Appli
3	1368	93.8	274	2	US-09-016-366A-21 Sequence 21, Appli
4	1368	93.8	274	2	US-08-978-404B-16 Sequence 16, Appli
5	1363	93.5	273	2	US-09-016-366A-19 Sequence 19, Appli
6	1363	93.5	267	2	US-08-978-404B-14 Sequence 14, Appli
7	1344	92.2	267	2	US-09-016-366A-23 Sequence 23, Appli
8	1344	92.2	245	4	US-08-978-404B-18 Sequence 18, Appli
9	1274	87.4	245	4	US-08-944-483-69 Sequence 69, Appli
10	1274	87.4	275	2	US-09-016-366A-17 Sequence 17, Appli
11	1274	87.4	275	2	US-08-978-404B-12 Sequence 12, Appli
12	1092	74.9	276	2	US-09-016-366A-15 Sequence 15, Appli

13	1092	74.9	276	2	US-08-978-404B-21
14	1075	73.7	270	2	US-08-978-404B-8
15	1057	72.5	273	2	US-08-978-404B-3
16	1055	72.4	274	2	US-08-978-404B-5
17	1039	71.3	273	2	US-08-978-404B-6
18	702	48.1	190	2	US-08-845-998-4
19	702	48.1	190	4	US-09-206-537-4
20	702	48.1	190	4	US-09-430-854-4
21	698	47.9	190	2	US-08-845-998-6
22	698	47.9	190	4	US-09-206-537-6
23	698	47.9	190	4	US-09-430-854-6
24	684.5	46.9	269	2	US-08-978-404B-10
25	563.5	38.6	290	4	US-09-386-653A-7
26	554.5	38.0	315	4	US-09-386-653A-9
27	507	34.8	319	4	US-09-386-652-12
28	507	34.8	328	4	US-09-386-652-11
29	506	34.7	299	4	US-08-944-483-66
30	501.5	34.4	317	4	US-09-386-652-7
31	494	33.9	314	4	US-09-008-271A-3
32	492.5	33.8	327	4	US-09-386-652-8
33	480.5	33.0	306	4	US-09-386-652-53
34	476	32.6	454	4	US-09-518-046-2
35	460.5	31.6	418	1	US-08-508-448C-25
36	460.5	31.6	418	4	US-09-370-838-82
37	460.5	31.6	418	4	US-09-370-838-83
38	460	31.6	238	4	US-08-944-483-64
39	458.5	31.4	418	4	US-09-370-838-62
40	449.5	30.8	232	1	US-08-508-448C-19
41	446	30.6	248	4	US-08-944-483-63
42	445	30.5	638	2	US-08-681-151-3
43	431.5	29.6	241	4	US-08-944-483-59
44	431	29.6	791	1	US-08-643-219-1
45	431	29.6	791	3	US-08-851-350-1

ALIGNMENTS

RESULT 1
US-09-079-970A-5
Sequence 5, Application US/09079970A
Patent No. 6274366
GENERAL INFORMATION:
APPLICANT: Mafilt, Mark A.
APPLICANT: Niles, Andrew L.
APPLICANT: Haak-Frendscho, Mary
TITLE OF INVENTION: Enzymatically-Active Recombinant Human
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: Intellectual Property Department
STREET: 8000 Excelsior Drive, Suite 401
CITY: Madison
STATE: WISCONSIN
COUNTRY: U.S.A.
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/079, 970A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Leone, Joseph T.
REGISTRATION NUMBER: 37,170
REFERENCE/DOCKET NUMBER: 34506, 073
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
INFORMATION FOR SEQ ID NO.: 5:
SEQUENCE CHARACTERISTICS:

Sequence 21, Appli
Sequence 8, Appli
Sequence 3, Appli
Sequence 5, Appli
Sequence 6, Appli
Sequence 4, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 10, Appli
Sequence 7, Appli
Sequence 12, Appli
Sequence 11, Appli
Sequence 66, Appli
Sequence 7, Appli
Sequence 3, Appli
Sequence 8, Appli
Sequence 53, Appli
Sequence 2, Appli
Sequence 25, Appli
Sequence 82, Appli
Sequence 83, Appli
Sequence 62, Appli
Sequence 19, Appli
Sequence 63, Appli
Sequence 3, Appli
Sequence 59, Appli
Sequence 1, Appli

LENGTH: 249 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-079-970A-5

Alignment Scores:

Pred. No.:	1-5e-104	Length:	249
Score:	1387.00	Matches:	248
Percent Similarity:	99.60%	Conservative:	0
Best Local Similarity:	99.60%	Mismatches:	1
Query Match:	95.13%	Indels:	0
DB:	4	Gaps:	0

US-09-598-982-20 (1-771) x US-09-079-970A-5 (1-249)

```

OY 7 CTCGAGAAAGANTGTCGGGGGTGAGAGGCCCCAGAGCAAGTGGCCCTGCAGCTG 66
Db 1 LeuGIuIySArgIleValGIyGIyGIuIaIaProArGserLySTRProIrpGIuVal 20
OY 67 AGCGTAGAGTCCAGCGCCCATCTGAGTGAACCTTGGCGGGGCTCCCTCATCCACCCC 126
Db 21 SerIeuArGValInHsGIyPrOTyTrpMetHisPheCysGIyGIySerIeuIleHisPro 40
OY 127 CAGTGGTGTGACCGCGCGCGTGCCTGGAGCCGAGCAAGTCAAGATCGGCCCTC 186
Db 41 GIuTrpValIleuThrAlaIleHisCysValGIyProArpValLyAspIleuAlaIleu 60
OY 187 AGGTGCAATGCGGGAGCAGACCTGTACTACAGAACCAAGCTCTCCGGTCAGCAGG 246
Db 61 ArgValGIuIeArGGIuGIuInHsIleuTyTrGIuAspGIuIeArGIuIeProValSerArg 80
OY 247 ATCATGTCGACCCCGACATCTGTACACCGCCAGATCGAGCGAGCAATGGCCCTGTGAG 306
Db 81 IleIleValInHsProGIuPheTyTrpThrAlaGIuIleGIyAlaAspIleIleIleuGIu 100
OY 307 CTGAGAGAGCGGTGAGAGTCTCCAGCCAGCTCCACAGCGTCCACCTGCCCTCA 366
Db 101 LeuGIuGIuProValIyValSerHisValInHsThrValThrIleuProIaIaSer 120
OY 367 CAGACCTTCCCCCGGAGATGCGGTGCGTCACTGGCTGGCGGAGATGGACAATGAT 426
Db 121 GIuTrpPheProGIuIyMetProCysTrpValThrGIyTrpGIyAspValAspAsnAsp 140
OY 427 GAGCGCCCTCCCGCCGCAATTCCTCTGAAGCAGGTGAAGTCCCATATGGAACACAC 486
Db 141 GIuArgIleProIProPheProIleuLyGIuValIyValProIleMetGIuAsnHis 160
OY 487 ATTTGTGACGCAAAATACACCTTGCGCGCTACACGGAGAGAGAGCTCCGATCGCGT 546
Db 161 IleCysAspAlaIyIyTrpHisIleuGIyAlaTyTrpGIyAspAspValAlaGIyAlaArg 180
OY 547 CAGCAGATGCTGTGTCGGGGAACACCGGAGGACTATGCGCAGGCGACTCCGAGGG 606
Db 181 AspAspMetIleCysAlaGIyAsnThrArgAspSerCysGIuGIyAspSerGIyGIy 200
OY 607 CCGCGGTGTGCAAGGTGAATGGACGCTGCGTGGCGAGCGGCGGCTGTAGCTGGGCGAG 666
Db 201 ProIleuValIyCysIyValAsnGIyThrTrpIleuGIuAlaGIyValIyValSerTrpIyGIu 220
OY 667 GAGTGTGCGCCAGCCCAACCGGCTGACATCTACACCGCTGTACCTACTTACTTGAGACTG 726
Db 221 GIyCysAlaGIuInProAsnArgProGIyIleTyTrpArgValThrTyTrpIleuAspTrp 240
OY 727 ATTCACCACTATGTCCCAAAAAGCCG 753
Db 241 IleHisIyValIyProIyLysPro 249

```

RESULT 2

US-09-079-970A-6
 ; Sequence 6, Application US/09079970A
 ; Patent No. 6274366
 ; GENERAL INFORMATION:

```

: APPLICANT: Maffitt, Mark A.
: APPLICANT: Niles, Andrew L.
: APPLICANT: Haak-Frendscho, Mary
: TITLE OF INVENTION: Enzymatically-Active Recombinant Human
: TITLE OF INVENTION: Beta-tryptase and Method of Making Same
: NUMBER OF SEQUENCES: 6
: CORRESPONDENCE ADDRESS:
: ADDRESS: Intellectual Property Department
: STREET: 8000 Excelsior Drive, Suite 401
: CITY: Madison
: STATE: WISCONSIN
: COUNTRY: U.S.A.
: ZIP: 53717-1914
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/079,970A
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Leone, Joseph T.
: REGISTRATION NUMBER: 37,170
: REFERENCE/DOCKET NUMBER: 34506.073
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (608) 831-2100
: TELEFAX: (608) 831-2106
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 245 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-079-970A-6

```

Alignment Scores:

Pred. No.:	5.12e-103	Length:	245
Score:	1368.00 <td>Matches:</td> <td>244 </td>	Matches:	244
Percent Similarity:	99.59% <td>Conservative:</td> <td>0 </td>	Conservative:	0
Best Local Similarity:	99.59% <td>Mismatches:</td> <td>1 </td>	Mismatches:	1
Query Match:	93.83% <td>Indels:</td> <td>0 </td>	Indels:	0
DB:	4 <td>Gaps:</td> <td>0 </td>	Gaps:	0

US-09-598-982-20 (1-771) x US-09-079-970A-6 (1-245)

```

OY 19 ATGTGGGGGGTCAGAGGCCCCAGAGCAAGTGGCCCTGCAGAGCTGAGACTC 78
Db 1 IleValGIyGIuGIuAlaIaProArGserLySTRProIrpGIuValSerIeuArGIyVal 20
OY 79 CAGCGCCCATACTGATGACACTTTCGGCGGGGCTCCCTCATCCACCGCCAGTGGTGTG 138
Db 21 HisGIyProTyTrpMetHisPheCysGIyGIySerIeuIleHisProGIuIrpValIleu 40
OY 139 ACCGCGCGCGGTGCGTGGAGCCGAGCTCAAGATCTGAGCGCCCTGAGGTCGAATG 198
Db 41 ThrAlaIaIaHisCysValGIyProAspValIyAspIleuAlaIaIeGIuIleGIuIlePro 60
OY 199 CGGAGAGCAGACCTTACTACAGAGCAGACTGTGCCGGTTCAGAGATCATCGTGCAC 258
Db 61 ArgGIuGIuInHsIleuTyTrGIuAspGIuIleuProValSerArgIleIleValHis 80
OY 259 CCACAGTTCTTACACCGCCAGATCGGAGCGGACATCGCCCTGTGAGCTGAGAGACCG 318
Db 81 ProGIuPheTyTrpAlaGIuIleGIyAlaAspIleAlaIeGIuIleGIuIleGIuIlePro 100
OY 319 GTGAGGTCTCCAGCCAGTCCACAGCGGACCCCTGGCCCTCAGAGACTTCCCC 378
Db 101 ValIyValSerSerHisValHisThrValThrIleuProIaIaSerGIuThrPhePro 120
OY 379 CCGGGATGCGCTGCTGGGTCACTGGCTGGGCGATGTGACAAATGATGAGCGCTCCCA 438

```

Db 121 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspLysIleuArgleuPro 140
QY 439 CCGCATTTCTCTGAGACAGGTGAAGTCCCATTAATGGAACACATTTGTGACGCA 498
Db 141 ProPheProIleuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 160
QY 499 AATACACACCTTGGCGCTACACGGAGAGACGTCGCCATCGCTCGTGACGACATGCTG 558
Db 161 LysTyrHisLeuGlyAlaLeuThrGlyAspAspValArgIleValAlaArgAspAspMetLeu 180
QY 559 TTGCGCGGGAACACCCGGAGGACTCATGCCAGGGCGACTCCGGAGGGCCCTGGTGTGC 618
Db 181 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 200
QY 619 AAGTGAATGGACCTGGCTGAGGGCGGCGTGCAGCTGGGGGAGGGCGTGGCCCGCAG 678
Db 201 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlnLysAlaGln 220
QY 679 CCCAACCGGCTGGCATCTACACCGGTGTACCTACTACTGACTGACTGATCCACCACTAT 738
Db 221 ProAsnArgProGlyIleLeuThrArgValThrTrpTyrLeuAspTrpIleHisHisTyr 240
QY 739 GTCCCAAAAACCG 753
Db 241 ValProLysLysPro 245

RESULT 3
US-09-016-366A-21
Sequence 21, Application US/09016366A
Patent No. 5955431
GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.
APPLICANT: Huang, Chifu
TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
TITLE OF INVENTION: INHIBITORS
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,366A
FILING DATE: January 30, 1998
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/037,090
FILING DATE: 05-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7093
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 274 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-016-366A-21

Alignment Scores: 5.25e-103 Length: 274
Pred. No.:

Score: 1368.00 Matches: 244
Percent Similarity: 99.59%
Best Local Similarity: 99.59%
Query Match: 93.83%
DB: 2 Gaps: 0

US-09-598-982-20 (1-771) x US-09-016-366A-21 (1-274)

QY 19 ATGTCGCGGGGTCAGAGGCGCCCGAGACAGTGGCCCTGGAGTGCAGTGCAGATC 78
Db 30 LLevalGlyGlyGlnAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 49
QY 79 CAGCGCCCATCTGATGACATTTGCGGGGGCTCCATCATCCACCCAGTGGGTGC 138
Db 50 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 69
QY 139 ACCGCGCGCGCTGCTGGAGCCGAGCTCAAGATCTGGCCCTCAGGTGCAACTG 198
Db 70 ThrAlaAlaHisCysValGlyProAspValLysAspLeuAlaLeuArgValGlnLeu 89
QY 199 CCGGAGCCACCACTCTACTACAGAGACCACTGCTGCCGTGAGAGATCATCTGTCAC 258
Db 90 ArgGlnGlnHisLeuTyrTrpGlnAspGlnLeuProValSerArgIleLeuValHis 109
QY 259 CCACAGTTCTACACCGCCAGATCGGAGCGGACATCGCCCTGAGCTGAGAGACCG 318
Db 110 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlnIleuGlnIleu 129
QY 319 GTGAAGTGTCCAGCCACGTCACACAGGTACCTGCCCTGCCCTCAGAGACTTTCCCG 378
Db 130 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGlnThrPhePro 149
QY 379 CCGGGATGCCCTGCTGGGTCACTGGCTGGGGCGATGTGAGCAATGATGAGGCGCTCCA 438
Db 150 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspLysIleuArgleuPro 169
QY 439 CCGCATTTCTCTGAGACAGGTGAAGTCCCATTAATGGAACACACATTTGTACGCA 498
Db 170 ProPheProIleuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 189
QY 499 AATACACACCTTGGCGCTACACGGAGAGACGTCGCCATCGCTCGTGACGACATGCTG 558
Db 190 LysTyrHisLeuGlyAlaLeuThrGlyAspAspValArgIleValAlaArgAspAspMetLeu 209
QY 559 TTGCGCGGGAACACCCGGAGGACTCATGCCAGGGCGACTCCGGAGGGCCCTGGTGTGC 618
Db 210 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 229
QY 619 AAGTGAATGGACCTGGCTGAGGGCGGCGTGCAGCTGGGGGAGGGCGTGGCCCGCAG 678
Db 230 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlnLysAlaGln 249
QY 679 CCCAACCGGCTGGCATCTACACCGGTGTACCTACTACTGACTGACTGATCCACCACTAT 738
Db 250 ProAsnArgProGlyIleLeuThrArgValThrTrpTyrLeuAspTrpIleHisHisTyr 269
QY 739 GTCCCAAAAACCG 753
Db 270 ValProLysLysPro 274

RESULT 4
US-08-978-404B-16
Sequence 16, Application US/08978404B
Patent No. 5968782
GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.
APPLICANT: Huang, Chifu
TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
TITLE OF INVENTION: FIBRINOGEN
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston

2
D.D. :
camps :

US-09-598-982-20 (1-771) x US-09-016-366A-19 (1-273)

```

OY 19 ATCGCGGGGTCAGAGAGGCCCCCAGAGCAAGTGGCCCTGGAGCTGAGAGTC 78
Db 29 ILevalGlyGlnGlnAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 48
OY 79 CACGAGCCATCTAGATGACCTTCTGCGGGGGCTCCCTCATGCAACCCCACTGGGTGCTG 138
Db 49 HisGlyProGlyTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 68
OY 139 ACCGCGCGGCGTGGTGGGACCCGAGCTCAAGATCTGGCCCCCTCAGGGTGCACATG 198
Db 69 ThrAlaIleHisCysValGlyProAspValLysAspLeuAlaIleValArgValGlnLeu 88
OY 199 CGGGAGCAGACCTCTACTACAGAGCAGCTGCTCGCGGCGCAGAGATGATGCTGCAC 258
Db 89 ArgGlnGlnHisLeuTrpGlnAspGlnLeuProValSerArgIleIleValHis 108
OY 259 CCACAGTTCTACACCGCCAGATCGAGCGGACATCGCCCTGCTGAGACTGGAGAGCGG 318
Db 109 ProGlnPheTrpTrpAlaGlnIleGlyAlaAspIleAlaLeuGlnLeuGlnLysPro 128
OY 319 GTGAAGGTCTCCAGCCACGCTCCACAGGTCACCTGCCCCCTGCTCAGAGACTTCCCC 378
Db 129 ValAsnValSerSerHisValIleHisThrValThrLeuProAlaSerGlnThrPhePro 148
OY 379 CCGGGAGATGGCGTGGTCTGCTGAGCTGGGGCGCATGTGGACATGATGAGCGCTGCCA 438
Db 149 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGlnArgLeuPro 168
OY 439 CCGCATTTCTCTGAGAGCAGGTGAAGTCCCATTAATGAAACACCATTTGTGACGCA 498
Db 169 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 188
OY 499 AAATACCACTTGGCGCTTACACGAGGAGACGAGTCCGATGCTCCGTGACAGATGCTG 558
Db 189 LysTyrHisLeuGlnGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 208
OY 559 TGTGCGCGGAGACACCGCGAGGAGCTCATGCGGAGGCGACTCCGAGAGGCGCCCTGCTGTC 618
Db 209 CysAlaGlyAsnTrpArgValAspSerCysGlnGlyAspSerGlyGlyProLeuValLys 228
OY 619 AAGGTGAATGACACTGCTGCTCAGAGCGGCGTGGTCACTGGGCGAGAGGCTGTGCCAC 678
Db 229 LysValAsnGlnGlyTrpLeuGlnAlaGlyValValSerTrpGlyGlnGlyCysAlaGln 248
OY 679 CCGAAGCGGCTGCGATCTACACCGGCTGTCACCTACTGACTGGATCGATCCACCATAT 738
Db 249 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisIleTyr 268
OY 739 GTCCCAAAAACGCG 753
Db 269 ValProLysLysPro 273

```

RESULT 6

US-08-978-404B-14

Sequence 14, Application US/08978404B

Patent No. 5968782

GENERAL INFORMATION:

APPLICANT: Stevens, Richard L.

TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES

TITLE OF INVENTION: FIBRINOGEN

NUMBER OF SEQUENCES: 74

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wolf, Greenfield & Sacks, P.C.

STREET: 600 Atlantic Avenue

CITY: Boston

STATE: MA

COUNTRY: U.S.A.

ZIP: 02210-2211

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/978,404B

FILING DATE: 25-NOV-97

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/032,354

FILING DATE: 04-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Plumer, Elizabeth R.

REGISTRATION NUMBER: 36,637

REFERENCE/DOCKET NUMBER: B0801/7090

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-720-3500

TELEFAX: 617-720-2441

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 273 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: No. 5968782e

US-08-978-404B-14

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

2

Gaps:

0

1.33e-102

1363.00

99.18%

93.48%

2

Gaps:

0

US-09-598-982-20 (1-771) x US-08-978-404B-14 (1-273)

```

OY 19 ATCGCGGGGTCAGAGAGGCCCCCAGAGCAAGTGGCCCTGGAGCTGAGAGTC 78
Db 29 ILevalGlyGlnGlnAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 48
OY 79 CACGAGCCATCTAGATGACCTTCTGCGGGGGCTCCCTCATGCAACCCCACTGGGTGCTG 138
Db 49 HisGlyProGlyTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 68
OY 139 ACCGCGCGGCGTGGTGGGACCCGAGCTCAAGATCTGGCCCCCTCAGGGTGCACATG 198
Db 69 ThrAlaIleHisCysValGlyProAspValLysAspLeuAlaIleValArgValGlnLeu 88
OY 199 CGGGAGCAGACCTCTACTACAGAGCAGCTGCTCGCGGCGCAGAGATGATGCTGCAC 258
Db 89 ArgGlnGlnHisLeuTrpGlnAspGlnLeuProValSerArgIleIleValHis 108
OY 259 CCACAGTTCTACACCGCCAGATCGAGCGGACATCGCCCTGCTGAGACTGGAGAGCGG 318
Db 109 ProGlnPheTrpTrpAlaGlnIleGlyAlaAspIleAlaLeuGlnLeuGlnLysPro 128
OY 319 GTGAAGGTCTCCAGCCACGCTCCACAGGTCACCTGCCCCCTGCTCAGAGACTTCCCC 378
Db 129 ValAsnValSerSerHisValIleHisThrValThrLeuProAlaSerGlnThrPhePro 148
OY 379 CCGGGAGATGGCGTGGTCTGCTGAGCTGGGGCGCATGTGGACATGATGAGCGCTGCCA 438
Db 149 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGlnArgLeuPro 168
OY 439 CCGCATTTCTCTGAGAGCAGGTGAAGTCCCATTAATGAAACACCATTTGTGACGCA 498
Db 169 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 188
OY 499 AAATACCACTTGGCGCTTACACGAGGAGACGAGTCCGATGCTCCGTGACAGATGCTG 558
Db 189 LysTyrHisLeuGlnGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 208

```

QY	559	TTGTCGGGGGAACACCCGGGAGGACATCAATGACCGGCGACATCCGAGGGCCCCCTGGGTGC	6
Db	209	CysAlaIaGlnThrArgArgArgAspSerCysGlnGlnAspSerGlyAlaGlyProLeuValCys	22
QY	619	AAGGTGAATGGCACCTGGCTGCAGCGGGCGGTGGTCACTGGGGCGAGGGCTGTGCCAG	67
Db	229	LysValAlaGlnGlyThrTrpLeuGlnIaGlyValValSerTrpGlyGlnGlyCysAlaGln	24
QY	679	CCCAACGGGCGCTGGCATCTACACCCGCTGACACTACTACTTGGAGCTGGATCCACCACTAT	73
Db	249	ProAsnArgProGlyIleThrThrArgValThrTyThrLeuAspTrpIleHisHisTy	26
QY	739	GTGCCCCAAAAGCCG	753
Db	269	ValProLysLysPro	273
RESULT 7			
		US-09-016-366A-23	
		Sequence 23, Application US/09016366A	
		Patent No. 5955431	
		GENERAL INFORMATION:	
		APPLICANT: Stevens, Richard L.	
		APPLICANT: Huang, Chifu	
		TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE	
		TITLE OF INVENTION: INHIBITORS	
		NUMBER OF SEQUENCES: 65	
		CORRESPONDENCE ADDRESS:	
		ADDRESSEE: Wolf, Greenfield & Sacks, P.C.	
		STREET: 600 Atlantic Avenue	
		CITY: Boston	
		STATE: MA	
		COUNTRY: U.S.A.	
		ZIP: 02210-2211	
		COMPUTER READABLE FORM:	
		MEDIUM TYPE: Diskette	
		COMPUTER: IBM Compatible	
		OPERATING SYSTEM: DOS	
		SOFTWARE: PASTSEQ for Windows Version 2.0	
		CURRENT APPLICATION DATA:	
		APPLICATION NUMBER: US/09/016,366A	
		FILING DATE: January 30, 1998	
		CLASSIFICATION: 530	
		PRIOR APPLICATION DATA:	
		APPLICATION NUMBER: 60/037,090	
		FILING DATE: 05-FEB-1997	
		ATTORNEY/AGENT INFORMATION:	
		NAME: Plumer, Elizabeth R.	
		REGISTRATION NUMBER: 36,637	
		REFERENCE/DOCKET NUMBER: B0801/7093	
		TELECOMMUNICATION INFORMATION:	
		TELEPHONE: 617-720-3500	
		TELEFAX: 617-720-2441	
		TELEX:	
		INFORMATION FOR SEQ ID NO: 23:	
		SEQUENCE CHARACTERISTICS:	
		LENGTH: 267 amino acids	
		TYPE: amino acid	
		STRANDEDNESS: single	
		TOPOLOGY: linear	
		US-09-016-366A-23	
Alignment Scores:			
		Pred. No.:	
		4,54e-101	267
		Score:	1344.00
		Percent Similarity:	98.37%
		Best Local Similarity:	98.37%
		Query Match:	92.16%
		DB:	2
		Gaps:	0
QY	19	ATGCTGGGGGTAGAGAGCCCCCAGAGCAAGTGGCCCTGGCAGGTGAACCTGAGAGTC	78

Db	23	ILEVALGIYGLINGLVALAIPARQASERTYSTRPRQTRGILVALSERLEUARGVAL	42
QY	79	CACGGCCCATKCTGGATGTCACATTTCTGGGGGGCTCCCATATCCACSSCCCACTGGGTCTG	133
Db	43	ARGASPARGTATRTTRPMETHNISPhECYSGLYGLYSERLEULENIIISPROGINTPTRVALL	62
QY	139	ACCGCGCGGGGTGGGTGGAGCCGGAGAGCTCAAGGTCMGCGCGCCCTCAGAGCTCACTG	198
Db	63	THIALAALANISCSVALGILPROASPRVALYLSASPLAUALAALALVALGILNLEU	82
QY	199	CGGGAGCAGACACCTCTACTACACAGAGACCAGCTGCTGCCGTACACAGATCATCTGCAC	255
Db	83	ARGILUGLNIHLSLEUTYTUGLNASPRLEUENUPROVALSERARGILELEVALHIS	104
QY	259	CCACAGTTCTACACCGCCCAATGGAGCGAGACATCGCCCTCTGTGAGACTGGAGAGCCG	311
Db	103	PROGINTPEHETRTHTHVALAGLNIHGLYALASPLIALLALEUENGLIULENGLIUPRO	122
QY	319	GTMAAGCTCTCCAGCGACCTCCACACGGTCAACGCTCCGCCCCCTGGCTCAGACACTTCCC	378
Db	123	VALLYSVALSERSEHISVALHISHTHVALTHLEUPROPROLALASERGLIUTHPhETRO	142
QY	379	CCGGGAGTCCGCTGGGTGGTCACTGGCGTGGGGCGCATGTGACAAATGATGAGCGCTCCA	433
Db	143	PROGLMETPROCTSTRTPTVALTHGILYTRPRILYASPRVALASPRASNAAPGILARGLEUPRO	162
QY	439	CCGGCATTTCTCTGTAGACACAGTGAAGTCCCCATATGAGAAACACATTTGTGACGCA	498
Db	163	PROPRORPHEPROLEULYSGILNVALLYSVALPROLLEMETGLIUAHNIHISLIECYSASPRAL	182
QY	499	AAATPACSCCTTTGGGGCTTACACGGGAGACAGACGTCGTCATCGTCCGAGACATGCTG	555
Db	183	LYSTYHNIHLSLEUGLYALATYTHGLYASPRASPRVALARGILEVALARGASPRMETLEU	207
QY	559	TGTGCCGGGAAACACCCGGAGGAGACTCATGCCAGGGGCACTCCGAGAGGGCCCTGATGTGC	618
Db	203	CYSALAGLYANHTHARGARGASPRSECTYSGILNLYASPRSEGLYGLIYPROLEUVALCYS	222
QY	619	AAGGTGAATGGACACTGGCTGCAGAGCGGGCTGTGACTGAGCTGGGCGAGGCTGTGCCAG	678
Db	223	LYSVALANGLYTHTRPLEUGLNIHALAGLYALVALISERTRGILYGLIYCYUALAGLN	242
QY	679	CCCAACGGCGCTGGCATGTACACCCGTCGTCAACACTCTTGAGACTGGATGCACACAT	733
Db	243	PROASNAARGPROGLIILETYTHTRARGVALTHRTYTYULEASPRTRIENHISHTYR	262
QY	739	GTCCCAAAAAGCCG 753	
Db	263	VALPROLYSLYSPRO 267	
RESULT 8			
US-08-978-404B-18			
: Sequence 18, Application US/08978404B			
: Patent No. 5968782			
GENERAL INFORMATION:			
APPLICANT: Stevens, Richard L.			
TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES			
TITLE OF INVENTION: FIBRINOGEN			
NUMBER OF SEQUENCES: 74			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Wolf, Greenfield & Sacks, P. C.			
STREET: 600 Atlantic Avenue			
CITY: Boston			
STATE: MA			
COUNTRY: U.S.A.			
ZIP: 02210-2211			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Diskette			
COMPUTER: IBM Compatible			
OPERATING SYSTEM: DOS			
SOFTWARE: FastSeq for Windows Version 2.0			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/978,404B			

FILING DATE: 25-NOV-97
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/032,354
FILING DATE: 04-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7090
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5968782e
US-08-978-404B-18

Alignment Scores:
Pred. No.: 4.54e-101 Length: 267
Score: 1344.00 Matches: 241
Percent Similarity: 98.37% Conservative: 0
Best Local Similarity: 98.37% Mismatches: 4
Query Match: 92.18% Indels: 0
DB: Gaps: 0

US-09-598-982-20 (1-771) x US-08-978-404B-18 (1-267)

QY 19 ATGCTGGGGGTCAGAGGCCCCAGAGAGTGGCCCTGGAGGTGACCTGAGAGTC 78
DB 23 TLevalGlyGlnGlnLarProAqSerLysTrpTrpGlnValSerLeuAqGVal 42
QY 79 CAGGSCCATGATGATGACTTCTGCGGGGCTCCCTCATCCACCCAGTGGTGTG 138
DB 43 AqdaRAqGlyTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 62
QY 139 ACCGCGCGCGCTGCTGGGAGCCGAGCTCAAGATCTGGCCGCGCTCAGGCTGCACTG 198
DB 63 ThrAlaIleHisCysValGlyProAspValLysAspLeuAlaLeuAqGValGlnLeu 82
QY 199 CGGGAGCAGCAGCTGACTACAGCAGCAGCTGCGCTGCGCAGCAGGATGATCGCGAC 258
DB 83 AqGlnGlnHisLeuTrpGlnAspGlnLeuProValSerTrpGlnIleValHis 102
QY 259 CCAGACTTTCAGACCCAGATCGAGCGGAGCATGCGCTGCGAGCTGGAGAGCGG 318
DB 103 ProGlnPheTrpThrAlaGlnIleGlyAlaAspIleAlaLeuGlnLeuGlnGlnPro 122
QY 319 GTGAGGCTTCCAGCAGCAGCTCAGACGCTGACCTGCGCTGCTCAGAGACTTCCCC 378
DB 123 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGlnTrpPhePro 142
QY 379 CGGGGATGCGCGTGGGCTGCTGCTGCGGCGAGTGGAGATGATAGCGCCCTCCCA 438
DB 143 ProGlyMetProCysTrpAlaThrGlyTrpGlyAspValAlaAspAsnAspLysLeuPro 162
QY 439 CGGCATTTCTCTGAAGCAGGTGAAGTCCCATTAATGAAGAAACACCATTTGTGACGCA 498
DB 163 ProProPheProLeuLysGlnValLysValProIleMetGlnAsnHisIleCysAspAla 182
QY 499 AATATCCACCTTGGCGCTACAGCGGAGAGAGAGTCCGATGCTCGTACAGACATGCTG 558
DB 183 LysTrpHisLeuGlyAlaTrpThrGlyAspAspValArgIleValArgAspAspMetLeu 202
QY 559 TGTGCGGGGAGAACCGCGGAGGAGCATATGCAAGGCGAGTCCGAGAGGCGCCCTGGTGTG 618
DB 203 CysAlaGlyAsnThrArgAspSerCysGlnGlyAspSerTrpGlyProLeuValCys 222
QY 619 AAGGTGAATGGACCTGCTGCTGAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 678

DB 223 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlnCysAlaGln 242
QY 679 CCCAACCGGCTCGCATCTACACCCGCTGACACTTACTTGATGATGATCAGACATAT 738
DB 243 ProAsnArgProGlyIleTrpThrArgValThrTrpLeuAspTrpIleHisIleTyr 262
QY 739 GTCCCAAAAGCCG 753
DB 263 ValProLysLysPro 267

RESULT 9
US-08-944-483-69
Sequence 69, Application US/08944483
Patent No. 6232456
GENERAL INFORMATION:
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GRAMADOS, EDWARD N.
APPLICANT: KLAAS, MICHAEL R.
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STEWART, KENT D.
APPLICANT: STROUPE, STEVEN D.
TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,483
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6183.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 245 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6232456e
US-08-944-483-69

Alignment Scores:
Pred. No.: 2.02e-95 Length: 245
Score: 1274.00 Matches: 226
Percent Similarity: 94.69% Conservative: 6
Best Local Similarity: 92.24% Mismatches: 13
Query Match: 87.38% Indels: 0
DB: Gaps: 0

US-09-598-982-20 (1-771) x US-08-944-483-69 (1-245)

|||||.....
Db 211 CysalaglYAsnSerGlnArgAspSerCysLysGlyAspSerGlyGlyProLeuValCys 230
QY 619 AAGGGAATGACCTGCTGAGCGGCGGTGTCAGTGGGGCGAGGCGTGTGCCAG 678
Db 231 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpAspGlnGlyCysAlaGln 250
QY 679 CCCAACCGCGCTGGCATGCACCGCGTGCACCTACTTACTGATGCATCCACCATAT 738
Db 251 ProAsnArgProGlyIleThrTrpArgValThrTrpLeuAspTrpIleHisHisTyr 270
QY 739 GTCCCAAAAAGCCG 753
Db 271 ValProLysLysPro 275
RESULT 11
US-08-978-404B-12
; Sequence 12, Application US/08978404B
; Patent No. 5968782
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
; TITLE OF INVENTION: FIBRINOGEN
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/978,404B
; FILING DATE: 25-NOV-97
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 60/032,354
; FILING DATE: 04-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE NUMBER: B0801/7090
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 275 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: NO. 5968782e
US-08-978-404B-12
Alignment Scores:
Pred. No.: 2,07e-95 Length: 275
Score: 1274.00 Matches: 226
Percent Similarity: 94.69% Conservative: 6
Best Local Similarity: 92.24% Mismatches: 13
Query Match: 87.38% Indels: 0
DB: 2 Gaps: 0
US-09-598-982-20 (1-771) x US-08-978-404B-12 (1-275)
QY 19 ATCTGGGGGGTTCAGAGAGCCCAAGAGAGAGAGTGGCGCTGAGAGCTGAGAGTC 78
Db 31 IieValGlyGlyGlnGlnAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 50

QY 79 CAGGGCCATAGTGCATCTTGGGGGGGCTCCCTATCCACCCAGTGGGTGCG 138
Db 51 ArgAspArgTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 70
QY 139 ACCGGCGCGCGCTGGCTGGAGCGGACGTCAAGGATCTGGCGCCCTCAGGGTGCAGTG 198
Db 71 ThrAlaAlaHisCysLeuGlyProAspValLysAspLeuAlaThrLeuArgValGlnLeu 90
QY 199 CGGAGACAGACCTCTACTACAGAGACAGCTGTGCGCGGTGCGGAGAGATCATCTGAC 258
Db 91 ArgGlnGlnHisLeuTyrTrpGlnAspGlnLeuProValSerArgIleIleValHis 110
QY 259 CCACAGTCTACACCGCCAGATCGAGCGGACATCCGCCCTCTGGAGCTGAGAGAGCGG 318
Db 111 ProGlnPheTyrIleIleGlnThrGlyAlaAspIleAlaLeuGlnGlyIleGlnGlyPro 130
QY 319 GTGAAGTCTCCAGCCACATCCACAGGTCACCGCTGCCCTGCTCAGAGACTTCCCG 378
Db 131 ValAsnIleSerSerArgValHisThrValMetLeuProProAlaSerGlnThrPhePro 150
QY 379 CCGGGAGTCCGTGCTGGCTACCTGGCTGGCGCATGTGACAATGATAGCGCTCCCA 438
Db 151 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGlnProLeuPro 170
QY 439 CCGCATTCCTCTGAGAGCGGTAAGGTCGCCCATATGGAACCAATTTGTAGACGA 498
Db 171 ProProPheProLeuLysGlnValLysValProIleMetGlnAsnHisIleCysAspAla 190
QY 499 AATACACACTTGGCGCTCTACAGCGAGACGAGTCCCATCGCTGCAGCAGCATGCTG 558
Db 191 LysTyrHisLeuGlnYAlaIleThrGlyAspAspValArgIleIleArgAspAspMetLeu 210
QY 559 TGTCCGGGGAACACCCGAGGACTGATGCCAGGGCGACTCCGAGAGGCCCTGTGTGC 618
Db 211 CysAlaGlyAsnSerGlnArgAspSerCysLysGlyAspSerGlyGlyProLeuValCys 230
QY 619 AAGGTGAATGCACCTGCTGAGCGGGCGGTGCAGCTGGGGCGAGCGGTGCCCGG 678
Db 231 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpAspGlnGlyCysAlaGln 250
QY 679 CCCAACCGCGCTGGCATGCACCGCGTGCACCTACTTACTGATGCATCCACCATAT 738
Db 251 ProAsnArgProGlyIleThrTrpArgValThrTrpLeuAspTrpIleHisHisTyr 270
QY 739 GTCCCAAAAAGCCG 753
Db 271 ValProLysLysPro 275
RESULT 12
US-09-016-366A-15
; Sequence 15, Application US/09016366A
; Patent No. 5955431
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; APPLICANT: Huang, Chifu
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,366A

```

1      FILING DATE: January 30, 1998
2      CLASSIFICATION: 530
3      PRIOR APPLICATION DATA:
4      APPLICATION NUMBER: 60/037,090
5      FILING DATE: 05-FEB-1997
6      ATTORNEY/AGENT INFORMATION:
7      NAME: Plumer, Elizabeth R.
8      REGISTRATION NUMBER: 36,637
9      REFERENCE/DOCKET NUMBER: BOB081/70933
10     TELECOMMUNICATION INFORMATION:
11     TELEPHONE: 617-720-3500
12     TELEFAX: 617-720-2441
13     TELEX:
14     INFORMATION FOR SEQ ID NO: 15:
15     SEQUENCE CHARACTERISTICS:
16     LENGTH: 276 amino acids
17     TYPE: amino acid
18     STRANDEDNESS: single
19     TOPOLOGY: linear
20     MOLECULE TYPE: protein
21     OS-09-016-366A-15

```

Alignment Scores:	
Pred. No.:	1 05e-80
Score:	1092.00
Percent Similarity:	84.80%
Best Local Similarity:	76.80%
Query Match:	74.90%
DB:	2
	Gaps: 1
	length: 276
	Matches: 192
	Conservative: 20
	Mismatches: 36
	Indels: 1

US-09-598-982-20 (1-771) x US-09-016-366A-15 (1-276)

0y	4	CCCTCCGAGAAACA-----ATCGTGGGGGTGAGGAGGCCCCACAGACAGTGGCC	57
		
Db	25	ProIaasngIvalGvalGlyIleValaIbLyIhIscIaIaSerGIuSerIyTtrPro	44
0y	58	TGGCAGTGCACCTGAGAGTCCACGGCCCAATACAGTGCACCTTTCGGGGGCTCCCTC	117
		
Db	45	TrpGIaIaSerIeuAaGpRheLyIeuaaInTyrIleIhISphecysGIuIySerIeu	64
0y	118	ATTCACCCCCACAGTGGGTCTGACCGCGCGCGCTGCTGGAGCCGACGTCAAGATCTG	177
		
Db	65	IlehISpGInTrrValIleuThrAlaIaIhIScysValGIyProhISileIySerPro	84
0y	178	GCCGCGCTCAGGGGCAACAGTGGGAGAGACCTTCATACCGAGACAGCTGCTCCG	237
		
Db	85	GInIeuPhehIyValGInIeuAaGpIuInTyrIleuTyrIyTrGIyAspGIuIeuIeSer	104
0y	238	GTACAGCAGATCATCTGTCACCCACAGTTCTACACCGCCAGATCGAGCGGACATGCC	297
		
Db	105	LeuAaAaGIIeValIaIhISProhISyTrIyThrIaIuGIuGIyAlaIaSpValaIa	124
0y	298	CTGCTGGAGCTGGAGAGCCGGTGAAGTCTTCACGCCACGTCCACCGGTACCCCTGCC	357
		
Db	125	LeuIeuGIuIeuGIuValProValaSnValSerThrIhISleIhISProIleSerIeuPro	144
0y	358	CCTGCTCAGAGACCTTCCCGGGGATGCGCGTGGGACATGCTGGGGGAGATG	417
		
Db	145	ProIaASerGIuThrPheProProGIyThISerIyStrIvalThrGIyTrpGIyAspIle	164
0y	418	GACAAATGATGAGCGCTCCACCGCCCATTTCTCTGAGAGCAGTGAAGGTCCCATATAG	477
		
Db	165	AspaAaAspGIuProIeuProProIyTrProIeuIySgIuValIyValProIleVal	184
0y	478	GAACACCAATTTGTGACGGCAAAATACCACCTTGGCGCTACACGGGAGACAGCTCCG	537
		
Db	185	GIuaISerIeuCysAspArGIyTyrIhISrGIyIleuTyrIhTrGIyAspAspPhePro	204
0y	538	ATCGTCCCTGACAGCATCTGTGTGGCGGAAACCCGGAGGAGCTATCCAGAGCGAC	597
		
Db	205	IleValIhISAspGIyMeIleuIeuCysAlaGIyISnThrIaGpAspSerIySgIuGIyAsp	224
0y	598	TCGCGAGGCGCCCTGTGTGCAAGGTGAATGGACCTGCTGCAGCGGCGTGTGTAGC	657

Db	225	SerGlyGlyProLeuValCysIstysValLysGlyThrProLeuGlnAlaGlyValValSer	244
Qy	658	TGGGCGAGGCGTGTGCCAGCCCAACGGCGCTGGCATCTACACCCGTCACCTACTAC	717
Db	245	TrpGlyGlnGlyCysAlaGlnProAsnLysProGlyIleTyrThrArgValThrTyr	264
Qy	718	TTGGACTGGATCCACCACTATGTCCCAAA	747
Db	265	LeuAspTrpIleHisArgTyrValProLeu	274

RESULT 13
US-08-978-404B-21

Alignment Scores:	
Pred. No.:	1.05e-80
Score:	1097.00
Percent Similarity:	84.80%
Best Local Similarity:	76.80%
Query Match:	74.90%
DB:	2
	Length: 277
	Matches: 19
	Conservative: 20
	Mismatches: 36
	Gaps: 1

US-09-598-982-20 (1-771) X US-08-978-404B-21 (1-276)

OY 4 CCCCTCGAGAAAGA-----ATGTCGGGGGTCAAGAGGCCACAGAGGAAGTGGCCC 57
 |||||
Dd 25 PROHlaasglnArgValGIylleIalEltyglrhysgnlhaIsarGeluseLlystPrrPro 44
 |||||
OY 58 TGGCAGGTAGCCCTAGAGACTCACGGCCCATCTGCATGACAATTCTGCGGGGGCTCCCTC 117
 |||||
Dd 45 TRPGInvalSerLeuAlrPhelystLeuaasnnyITrrlleisnphicSGIGLyISerieu 64

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 31, 2003, 05:24:20 ; Search time 46 Seconds

(without alignments)
4466.793 Million cell updates/sec

Title: US-09-598-982-20

Perfect score: 1458

Sequence: 1 gggccctcagagaagaat.....cgtgaagcgcgcgtcgt 771

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 segs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_n2p_model -DBY=x1h
-Q/cg22_1/USPTO_SPOOL/US09598982/rnat_27012003_073715_19505/app_query.fasta_1.967
-DB-A=Geneseq_101002 -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdt
-LIST=45 -DOCLITS=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=PCT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09598982.ecgn.1.1.0.0.rnat_27012003_073715_19505 -NCPU=6 -ICPU=3
-NO_XLIFY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A.Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1393	95.5	249	AAU12009	Human beta-II tryp
2	1388	95.2	249	AAU12017	Human beta-II tryp
3	1387	95.1	249	AAU55011	Human beta-tryptas
4	1387	95.1	249	AAU12007	Recombinant human
5	1384	94.9	249	AAU12011	Human beta-II tryp
6	1384	94.9	249	AAU12012	Human beta-II tryp
7	1382	94.8	249	AAU12006	Human beta-I tryp
8	1379	94.6	249	AAU12010	Human beta-II tryp
9	1379	94.6	249	AAU12019	Human beta-II tryp
10	1379	94.6	249	AAU12020	Human beta-II tryp
11	1374	94.2	245	AAU12013	Human beta-II tryp
12	1374	94.2	249	AAU12018	Human beta-II tryp
13	1369	93.9	245	AAU12021	Human beta-II tryp
14	1368	93.8	245	AAU12008	Recombinant human
15	1368	93.8	274	AAU64240	Human mast cell tr
16	1368	93.8	274	AAU63175	Human mast cell tr
17	1365	93.6	245	AAU12015	Human beta-II tryp
18	1365	93.6	245	AAU12016	Human beta-II tryp
19	1363	93.5	245	AAU55010	Human beta-tryptas
20	1363	93.5	245	AAU12005	Human mature beta-
21	1363	93.5	273	AAU64238	Human mast cell tr
22	1363	93.5	273	AAU63174	Human mast cell tr
23	1361	93.3	244	AAU25925	Human lung mast ce
24	1360	93.3	245	AAU12014	Human beta-II tryp
25	1360	93.3	245	AAU12023	Human beta-II tryp
26	1360	93.3	245	AAU12024	Human beta-II tryp
27	1355	92.9	245	AAU12022	Human beta-II tryp
28	1344	92.2	267	AAU64241	Human mast cell tr
29	1344	92.2	267	AAU63176	Human mast cell tr
30	1344	92.2	267	AAU64360	Protein TP52 diffe
31	1329	91.2	691	AAU14348	Human protease PRT
32	1274	87.4	275	AAU64237	Human mast cell tr
33	1274	87.4	275	AAU63173	Human mast cell tr
34	1092	74.9	276	AAU64242	Murine mast cell p
35	1092	74.9	276	AAU63172	Murine mast cell p
36	1077	73.9	276	ABU05437	Mouse mast cell pr
37	1075	73.7	270	AAU64239	Human mast cell pr
38	1057	72.5	273	AAU64233	Gerbil homologue o
39	1057	72.5	310	AAU14343	Murine mCP-7 zymo
40	1055	72.4	274	AAU64234	Human protease PRT
41	1039	71.3	273	AAU64235	Rat homologue of m
42	1006	69.0	273	AAU81826	Pig lung protease.
43	702	48.1	190	AAU68539	Amino acid sequenc
44	698	47.9	190	AAU68540	Amino acid sequenc
45	684.5	46.9	269	AAU64236	Canine homologue o

ALIGNMENTS

RESULT 1
AAU12009 standard; Protein: 249 AA.
AAU12009;
09-APR-2002 (first entry)
Human beta-II tryptase active site mutant H44A #1.
Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
enzyme; mutant; mutein.
Homo sapiens.
Synthetic.
WO200198470-A2.
27-DEC-2001.

PF	20-JUN-2001:	200IMO-US19681.	
XX			
XX	21-JUN-2000:	200OUS-0598982.	
XX			
PA	(PROM-)	PROMEGA CORP.	
XX			
PI	Maffit M,	Niles AL,	Haak-Frendscho M;
XX	WPI:	2002-114578/15.	
DR	N-PSDB:	AAS20775.	
XX			
PT	DNA construct for producing enzymatically-inactive proteolytic		
PT	trypsin, comprises DNA sequence encoding proteolytic trypsin having		
PT	an active site mutation -		
XX			
PS	Claim 8:	Page 84-85;	126pp; English.
XX			
CC	The present invention relates to recombinant human proteolytic		
CC	trypsin, active site mutants of these trypsin and the methods for		
CC	producing these. The method involves the production of a DNA expression		
CC	construct comprising a promoter operably linked to a secretion signal		
CC	sequence which is operably linked to a DNA sequence encoding a		
CC	proteolytic trypsin with an active site mutation (the construct drives		
CC	expression of a mature proteolytic trypsin that lacks enzymatic activity		
CC	due to the active site mutation, in hosts transformed to contain the		
CC	construct). The method is useful for producing enzymatically-active		
CC	beta-II trypsin. The active site mutants of proteolytic trypsin provide		
CC	a tool to investigate the structural and functional properties of the		
CC	protease and its enzymatic activity, and for modelling studies. The		
CC	enzymatically-active, recombinant proteolytic trypsin produced are		
CC	useful as an antigen to generate anti-human trypsin antibodies		
CC	and in drug screening for compounds which act as trypsin inhibitors,		
CC	antagonists, agonists, etc. AAU12009-AAU12024 represent recombinant		
CC	human beta-II trypsin active site mutants.		
XX			
SO	Sequence	249 AA:	
Alignment Scores:			
Prod. NO.:	1.43e-97	Length:	249
Score:	1393.00	Matches:	249
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	95.54%	Indels:	0
DB:	23	Gaps:	0
US-09-598-982-20 (1-771) x AAU12009 (1-249)			
OY	7	CTCGAAGAAAGATGTCGGGGGCTCAGAGAGGCCCGGAGCAATGCGCTGGCAGTG	66
Db	1	LeuGluIuYsArgIleValIglYglGlnIalAProArgSerIuYrProItrpGlnVal	20
OY	67	AGCGTGAAGTCCAGCGGCCCATCTAGATGCACTTCTCGGGGGCTCCCTCAATCCACCC	126
Db	21	SerLeuAArgValIhISglYrProIuYrTrpMenIhISPhcysglYlYserLeuIleIhISPro	40
OY	127	CAGTGGGCTGACCGCGCGCGCGCTGGTGGAGCCGAGACATGATGCGCGCGCTC	186
Db	41	GlnItrpValIleuThrAlaIalAalAcysValIglYrProAspValIylsAspIleuAalaleu	60
OY	187	AGGGTCAACTGCGGGGAGCAGACACCTTACTACACAGGACACAGCTGCTGGGTGAGCAG	246
Db	61	ArgValIglIleuAArgIuGlnIhISleuTYrTglAspIleuIleuIleuProValSerIuG	80
OY	247	ATCATGTGACCCCAACAGTGTACACCGGCCCAATGCGAGCGACATGCGCTGCTGGAG	306
Db	81	IleIleValIhISrProGlnIhISrProIuYrThrAlaGlnIleIeIYAlasrIleAaleuIleuGln	100
OY	307	CTGGAGGAGCGGGGAGAGGTCTCAGCAGCAGTCAACAGGCTACCTGGCGCCCTGGCTCA	366
Db	101	LeuGlnIuIleuProValIylsValSerSerIhISValIhISrValIhISrProIalSer	120
OY	367	GAGACCTTCCCCCGGGGATGCGCTGCTGGGTACTGGGTGGGGCGCATGTGACATGAT	426

Db	121	GLUTININPROPERIOGLYMETPROCYSTRVALTHNGLYTRPGLYSPVALASPASN	140
OY	427	GAGCGCCTCCACCCGCAATTCTCTGGAAGCAGGTGAGGTCCCAATATGCAAAACAC	486
Db	141	GUATGLeuPROPROphEProLeuLysGlnValLysValProIleMetGluASNHis	160
OY	487	ATTTCGTACGCAAAATACACACTTGCGCGCTCACGGGAGAGCAGTCGCGATCGTCCT	546
Db	161	ILeCYsASPAlaLysTYrHisLSeuGlyAlaTYrThGlyASPaspAlaArgIleValArg	180
OY	547	GACGACATCTGTGTGCCCGGAACACCCGAGGAGACTCATGCCAGGGCGACTCCGAGGG	606
Db	181	ASPASPMeLLeuCysAlaGlyASNThrArgArgASPserCysInGlyASPserGlyGly	200
OY	607	CCCCGTGTGTGCAAGCTGAATGGACCTTGCTGACGCGGCGCTGTCACTGGGCGAG	666
Db	201	ProLeuValCYsLysValASNGLYThrTRPLeuGlnAlaGlyValIleSerTRPGLYGLu	220
OY	667	GGCTGTGGCCACCCCAACCGGCTGGCATCTACACCCGTGTACCTACTCTTGAGACTGG	726
Db	221	GLYCysAlaGlnProASNArgProGlyIleTYrThrArgValThrTYrLLeuASPTrp	240
OY	727	ATCCACCACTATGTCCCCAAAAGCCG	753
Db	241	ILeHisHisTYrValProLysLysPro	249
RESULT 2			
AAU12017			
XX	AAU12017	standard: Protein: 249 AA.	
XX	AAU12017;		
XX	09-APR-2002	(first entry)	
DE	Human beta-II tryptase active site mutant H44A #3.		
XX	Human: proteolytic tryptase: protease; recombinant beta-II tryptase;		
KW	enzyme; mutant; mutein.		
XX	Homo sapiens.		
OS	Synthetic.		
XX	WO200198470-A2.		
PN	27-DEC-2001.		
PD	20-JUN-2001: 2001WO-US19681.		
XX	21-JUN-2000: 2000US-0598982.		
PR	(PROM-) PROMEGA CORP.		
XX	Maffit M, Niles AL, Haak-Frendscho M;		
XX	WPI: 2002-114578/15.		
DR	N-PSDB: AAS20783.		
XX	DNA construct for producing enzymatically-inactive proteolytic		
PT	tryptase, comprises DNA sequence encoding proteolytic tryptase having		
XX	an active site mutation -		
PS	Claim 8: Page 105-106; 126pp: English.		
XX	The present invention relates to recombinant human proteolytic		
CC	tryptases, active site mutants of these tryptases and the methods for		
CC	producing these. The method involves the production of a DNA expression		
CC	construct comprising a promoter operably linked to a secretion signal		
CC	sequence which is operably linked to a DNA sequence encoding a		
CC	proteolytic tryptase with an active site mutation (the construct drives		
CC	expression of a mature proteolytic tryptase that lacks enzymatic activity		
CC	due to the active site mutation. In hosts transformed to contain the		
CC	construct) the method is useful for producing enzymatically-active		

CC beta-II tryptase. The active site mutants of proteolytic tryptase provide
CC a tool to investigate the structural and functional properties of the
CC protease and its enzymatic activity, and for modelling studies. The
CC enzymatically-active, recombinant proteolytic tryptase produced are
CC useful as an antigen to generate anti-human tryptase antibodies
CC and in drug screening for compounds which act as tryptase inhibitors,
CC antagonists, agonists, etc. AAU12009-AAU12024 represent recombinant
CC human beta-II tryptase active site mutants.

XX Sequence 249 AA:

Alignment Scores:
Pred. No.: 3,41e-97 Length: 249
Score: 1388.00 Matches: 248
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 95.20% Indels: 0
DB: 23 Gaps: 0

US-09-598-982-20 (1-771) x AAU12017 (1-249)

OY 7 CTCGAGAAAGATCGTCGGGGTCAGAGAGCCCGCAGAGCAAGTGGCCCTGGCAGTG 66
Db 1 LeuGluValArgIleValIleGlyGlnGlnAlaProArgSerLysTrpProTrpGlnVal 20
OY 67 ACCCTGAGAGTCACAGCCCATCTACTGATGACATCTCTGGGGGGCTCCCTCATCCACCCC 126
Db 21 SerLeuArgValHisGlyProTrpTrpMetHisPheCysGlyLysSerLeuIleHisPro 40
OY 127 CAGTGGGTCCTAGCCCGCGCGTGGAGCCGAGCAGCATGAGATCTGGCCGCCCTC 186
Db 41 GlnTrpValLeuThrAlaIleAlaCysValIleGlyProAspValLysAspLeuAlaIleLeu 60
OY 187 AGGTCACACTCGCGGAGACAGACACCTCTACTACAGAGACAGCTGCTGCCGCTCAGCAG 246
Db 61 ArgValGlnLeuArgIleGlnHisLeuTrpTyrGlnAspIleLeuLeuProValSerArg 80
OY 247 ATCATCGTCACACCATCTTACACCGCCGAGATCGAGCGGAGCATCGCCCTGGCGAG 306
Db 81 IleIleValHisProGlnPheTrpThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGln 100
OY 307 CTGGAGAGCCGGTGAAGTCTTCACAGCAGCTCCACAGCGTCACCCCTGCCCTGCTCA 366
Db 101 LeuGlnGlnProValAsnValSerSerHisValHisThrValThrLeuProProAlaSer 120
OY 367 GAGACCTTCCCCCGGGGATGCCGCTGCTGGTCTGCTGGGGGCATGTGGACATGAT 426
Db 121 GlnThrPheProGlyMetProCysTrpValThrGlyTrpLeuLysPalaAspAsnAsp 140
OY 427 GAGCGCTCCACCGCATTTCTCTGAAGCAGGTGAAGTCCCGCATATGAAACACAG 486
Db 141 GlnArgLeuProProPheProPheProLeuLysGlnValLysValProIleMetGlnAsnHis 160
OY 487 ATTTGTGAGCAAAATACACCTTGCGCCCTACACGGAGACAGCAGCTCCGATCGTCCG 546
Db 161 IleCysAspAlaLysTrpHisLeuGlyAlaTyrThrGlyAspAspAlaArgIleValArg 180
OY 547 GAGAGACTGTCGTCCGGGAAACACCGGAGGAGACTCATGCCAGGGCAGCTCCGAGGG 606
Db 181 AspAspMetLeuCysAlaGlyAsnThrArgAspSerCysGlnIleLysPseArgIleGly 200
OY 607 CCCCTGGTGTGCAAGGTGAATGACACCTGGCTGACAGCGGGCGTGTCAAGTGGGCGAG 666
Db 201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValAspTrpGlyGln 220
OY 667 GCGTGTGCCAGCCCAACCGCGCTGGCATCTACACCGGTGTCACCTACTGACTGGACTGG 726
Db 221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTrpLeuAspTrp 240
OY 727 ATCCACCACTATGTCCCAAAAAGCCG 753
Db 241 IleHisIstIstValProLysLysPro 249

RESULT 3
AAV55011
ID AAV55011 standard. Protein: 249 AA.

XX AAV55011;

DT 18-FEB-2000 (first entry)

DE Human beta-tryptase protein sequence.

KW Beta-tryptase; human; DNA expression construct; protein production;
KW combinatorial library screening; X ray crystallography; antigen;
KW antibody generation.

OS Homo sapiens.

PN WO9960139-A1.

PD 25-NOV-1999.

PF 29-OCT-1998; 98WO-US22994.

PR 15-MAY-1998; 98US-0079970.

PA (PROM-) PROMEGA CORP.

PI Maffitt MA.; Niles AL, Haak-Frendscho M;

DR WPI; 2000-053300/04.

DR N-PSDB; AA240175.

PT New DNA expression construct for production of enzymatically active

PT recombinant human beta-tryptase -

PS Disclosure; Page 43-44; 50pp; English.

CC This sequence is the human beta-tryptase. The invention relates to a
CC DNA expression construct comprising (5' to 3') a promoter linked to a
CC signal sequence which is linked to a sequence encoding human
CC beta-tryptase. The DNA construct is useful for transforming host cells to
CC express, post translationally process and secrete enzymatically active
CC human tryptase. The method is useful for the production of large amounts
CC of tryptase with defined specifications. The transformant is useful for
CC pharmacological studies, combinatorial library screens and X ray
CC crystallographic studies. The tryptase produced allows for the
CC development of tryptase agonists and/or antagonists, is useful as an
CC antigen to generate antihuman tryptase antibodies in various animals,
CC can be used in screening for compounds which act as tryptase inhibitors,
CC antagonists, agonists etc. and to assay for the presence of tryptase in
CC biological or other solutions. Tryptase inhibitors, antagonists, agonists
CC etc. may be useful as therapeutics. The tryptase does not require any
CC post-expression or post-purification modifications or manipulations to
CC initiate tryptase activity and it has enzymatic activity which comprises
CC favourably with cadaveric tryptase. The availability of enzymatically
CC active tryptase facilitates the large scale screening of combinatorial
CC libraries for specific tryptase inhibitors as potential therapeutics and
CC advances the understanding of the biological significance of tryptase in
CC mast cell mediated diseases. The tryptase can be used to detect low
CC levels of tryptase.

XX Sequence 249 AA:

Alignment Scores:
Pred. No.: 4,06e-97 Length: 249
Score: 1387.00 Matches: 248
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 95.13% Indels: 0
DB: 21 Gaps: 0

US-09-598-982-20 (1-771) x AAV55011 (1-249)

OY 7 CTCGAGAAAGATCGTCGGGGTCAGAGAGCCCGCAGAGCAAGTGGCCCTGGCAGTG 66

```

Db      1 LeuGIuLysArgIleValGIgLyGInGInuaIaProArGSerLystrProtrpGInVal 20
OY      67 AGCCGAGAGTCCAGCGCCCATACAGTGCATCTGTGGGGGGCTCCCTCATCCAGCCC 126
Db      21 SerLeuArgValaHIsGlyProIyrrTrpMetHIsPheCySgIyGIySerLeuIleHIsPro 40
OY      127 CAGTGGGTGCTACCGCGCGCGGTGCGGTGGAGCCGAGGTCAAGATGTGGCCGGCTC 186
Db      41 GIuTrpValIleuThrAlaAlaHIsCySValGIyProAspValIlyAspLeuAlaIaleu 60
OY      187 AGGGTCAACTCGCGGAGCAGCAGCCTTACTACAGAGCAGCAGCTGCTGCGGTCAAGCAG 246
Db      61 ArgValGIuLeuArgGIuGIuGIuHIsLeuIyrrTyrgInaSPGIuLeuLeuProValSerArg 80
OY      247 ATCATCGTGCACCCACAGTTCATACACCGCCGAGTCCGAGCGGACATGCCCTGTGGAG 306
Db      81 IleIleValHIsProGIuPheTyrrThrAlaGInIleGIyAlaAspIleAlaLeuLeuGIu 100
OY      307 CTGAGAGACCGCGTGAAGGTCTCCAGCCAGCTCCACAGCGTCAACCTGCGCCCTGCTCA 366
Db      101 LeuGIuGIuProValIlyValSerSerHIsValHIsThrValIThrLeuProProAlaSer 120
OY      367 GAGACCTTCCCGCGGGGATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 426
Db      121 GIuTrpPheProProGIyMetProCystrpValIThrGIyTrpGIyAspValaSPaSPaSP 140
OY      427 GAGCGCCCTCCCGCCCATCTCTGTGAAGCAGGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 486
Db      141 GIuHArgLeuProProProPheProLeuLysGIuValIlyValIProIleMetGIuSPaHIs 160
OY      487 ATTGTGAGCGCAAAATACACCTTGCGCGCTGACAGCGGAGAGAGAGTCCGATCGTCGT 546
Db      161 IleCySAspAlaIlySerIyrrHIsLeuGIyAlaIyrrHnGIyAspAspAlaArgIleValaArg 180
OY      547 GAGGACATGCTGTGTGCGGGGAGAACACCGGAGGAGACTATGCCAGGCGAGTCCGAGGG 606
Db      181 AspAspMetLeuCysAlaGIyAsnThrArgAspSerCySgIyGIySPaSerGIyGIy 200
OY      607 CCCCGTGTGCAAGGTGAATGACACCTGTCGACGCGCGGGGTGCTGACGTGGCGGCGAG 666
Db      201 ProLeuValaCySlyValaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPa 220
OY      667 GGCTGTGCCACGCGCGCGCTGACATCTACACCGGTGCTGCTGCTGCTGCTGCTGCTG 726
Db      221 GIyCySAlaGInProaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPaSPa 240
OY      727 ATCCACCACTATGTCGCCCAAAAAGCCG 753
Db      241 IleHIsHIsTyrrValProIyLysPro 249
RESULT 4
AAU12007
ID      AAU12007 standard; Protein: 249 AA.
XX
XX      AAU12007;
XX
XX      09-APR-2002 (first entry)
XX
XX      Recombinant human beta-II tryptase.
DE
XX
XX      Human: proteolytic tryptase; protease; recombinant beta-II tryptase;
KW      enzyme; mutein; mutant.
XX
XX      Homo sapiens.
OS      Synthetic.
XX
XX      MO200198470-A2.
XX
XX      27-DEC-2001.
XX
XX      20-JUN-2001; 2001WO-US19681.
XX

```

```

PR      21-JUN-2000; 2000US-0598982.
XX
XX      (PROM-) PROMEGA CORP.
PA
XX
XX      Maffit M, Niles AL, Haak-Frendscho M;
PI
XX
XX      WPI: 2002-114578/15.
DR      N-PSDB: AAS20765.
XX
XX
XX      DNA construct for producing enzymatically-inactive proteolytic
PT      tryptase, comprises DNA sequence encoding proteolytic tryptase having
PR      an active site mutation -
XX
XX      Claim 60; Page 77-78; 126pp; English.
XX
XX      The present invention relates to recombinant human proteolytic
CC      tryptases, active site mutants of these tryptases and the methods for
CC      producing these. The method involves the production of a DNA expression
CC      construct comprising a promoter operably linked to a secretion signal
CC      sequence which is operably linked to a DNA sequence encoding a
CC      proteolytic tryptase with an active site mutation (the construct drives
CC      expression of a mature proteolytic tryptase that lacks enzymatic activity
CC      due to the active site mutation, in hosts transformed to contain the
CC      construct). The method is useful for producing enzymatically-inactive
CC      beta-II tryptase. The active site mutants of proteolytic tryptase provide
CC      a tool to investigate the structural and functional properties of the
CC      protease and its enzymatic activity, and for modelling studies. The
CC      enzymatically-active, recombinant proteolytic tryptase produced are
CC      useful as an antigen to generate anti-human tryptase antibodies
CC      and in drug screening for compounds which act as tryptase inhibitors,
CC      antagonists, agonists, etc. The present sequence represents recombinant
CC      human beta-II tryptase.
XX
XX      Sequence 249 AA:
SQ
XX
XX
XX      Alignment Scores:
XX      Pred. No.: 4,06e-97 Length: 249
XX      Score: 1387.00 Matches: 248
XX      Percent Similarity: 99.60% Conservative: 0
XX      Best Local Similarity: 99.60% Mismatches: 1
XX      Query Match: 95.13% Indels: 0
XX      DB: 23 Gaps: 0
XX
XX      US-09-598-982-20 (1-771) x AAU12007 (1-249)
OY      7 CTGAGAAAGATGTCGGGGGTGTCAGAGCGCCCGAGAGAGTGGCTGCGAGTGC 66
Db      1 LeuGIuLysArgIleValGIgLyGInGInuaIaProArGSerLystrProtrpGInVal 20
OY      67 AGCCGAGAGTCCAGCGCCCATACAGTGCATCTGTGGGGGGCTCCCTCATCCAGCCC 126
Db      21 SerLeuArgValaHIsGlyProIyrrTrpMetHIsPheCySgIyGIySerLeuIleHIsPro 40
OY      127 CAGTGGGTGCTACCGCGCGCGGTGCGGTGGAGCCGAGGTCAAGATGTGGCCGGCTC 186
Db      41 GIuTrpValIleuThrAlaAlaHIsCySValGIyProAspValIlyAspLeuAlaIaleu 60
OY      187 AGGGTCAACTCGCGGAGCAGCAGCCTTACTACAGAGCAGCAGCTGCTGCGGTCAAGCAG 246
Db      61 ArgValGIuLeuArgGIuGIuGIuHIsLeuIyrrTyrgInaSPGIuLeuLeuProValSerArg 80
OY      247 ATCATCGTGCACCCACAGTTCATACACCGCCGAGTCCGAGCGGACATGCCCTGTGGAG 306
Db      81 IleIleValHIsProGIuPheTyrrThrAlaGInIleGIyAlaAspIleAlaLeuLeuGIu 100
OY      307 CTGAGAGACCGCGTGAAGGTCTCCAGCCAGCTCCACAGCGTCAACCTGCGCCCTGCTCA 366
Db      101 LeuGIuGIuProValIlyValSerSerHIsValHIsThrValIThrLeuProProAlaSer 120
OY      367 GAGACCTTCCCGCGGGGATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 426
Db      121 GIuTrpPheProProGIyMetProCystrpValIThrGIyTrpGIyAspValaSPaSPaSP 140

```

QY 427 GAGCGCTCCACCGCATTTCTCTGAAGAGGTGAAGTCCCATATGAAAAACAC 486
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 141 GIUARGLEUPROPROBHEPROLEULYSGLINVALYVALPROLLEMETGLUASNHS 160
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 487 ATTTGTGACGCAAAATACCACTTGCGCGCTACAGGGAGAGAGAGCGCATCGTCCGT 546
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 161 IIECYSASPALALYSTYRTHISLEUGLYALATYRTHGLYASBPSPVALARGILEVALARG 180
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 547 GAGCAATCTGTGTGCCGGGAACACCGGAGGAGACTGTGCGAGGGCGACATCCGAGGG 606
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 ASPASPMELEUCYSALAGLYASNTHARGARGASPSETCYSGINGLYASPERGLYGLY 200
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 607 CCCCTGTGTGCAAGGTGAATGCACCTGGCTGCAGGCGGGCGGTGCTGAGTGGGCGAG 666
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 201 PROLEUVALCYSLYSVALASNGLYRTHRTPLEUGLINALAGLYVALVALSERTRPGLYGLU 220
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 667 GGCTGTGCCACCCCAACCGGCTGGCATCTACACCCGTGCTACCTACTGACTTGGACTGG 726
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 221 GLYCYSALAGLNPROMASNPTRPROGLYIETRYRTHARGVALTHRTYRTRLEUASPTRP 240
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 727 ATCCACCACTATGTCCCAAAAAAGCG 753
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 241 IIEHISHISTYRVALPROLYSLYSPRO 249
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 5
 AAU12011
 ID AAU12011 standard; Protein: 249 AA.
 XX AAU12011;
 DT 09-APR-2002 (first entry)
 DE Human beta-II tryptase active site mutant S194A #1.
 XX
 KM Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
 KM enzyme; mutant; mutein.
 XX
 OS Homo sapiens.
 OS
 XX WO200198470-A2.
 PN
 XX
 PD 27-DEC-2001.
 PD
 XX 20-JUN-2001; 2001WO-US19681.
 PF
 XX 21-JUN-2000; 2000US-0598982.
 PR
 XX (PROM-) PROMEGA CORP.
 PA
 PI Maffit M, Niles AL, Heak-Frendscho M;
 XX
 DR WPI: 2002-114578/15.
 DR
 DR N-PSDB: AAS20777.
 PT
 PT DNA construct for producing enzymatically-inactive proteolytic
 PT tryptase, comprises DNA sequence encoding proteolytic tryptase having
 PT an active site mutation -
 XX
 XX Claim 8, Page 90-91; 126pp; English.
 PS
 XX
 CC The present invention relates to recombinant human proteolytic
 CC tryptases, active site mutants of these tryptases and the methods for
 CC producing these. The method involves the production of a DNA expression
 CC construct comprising a promoter operably linked to a secretion signal
 CC sequence which is operably linked to a DNA sequence encoding a
 CC proteolytic tryptase with an active site mutation (the construct drives
 CC expression of a mature proteolytic tryptase that lacks enzymatic activity
 CC due to the active site mutation, in hosts transformed to contain the
 CC construct). The method is useful for producing enzymatically-active
 CC beta-II tryptase. The active site mutants of proteolytic tryptase provide
 CC a tool to investigate the structural and functional properties of the
 CC protease and its enzymatic activity, and for modelling studies. The

CC enzymatically-active, recombinant proteolytic tryptase produced are
 CC useful as an antigen to generate anti-human tryptase antibodies
 CC and in drug screening for compounds which act as tryptase inhibitors,
 CC antagonists, agonists, etc. AAU12009-AAU12024 represent recombinant
 CC human beta-II tryptase active site mutants.
 CC
 XX
 SQ Sequence 249 AA;
 Alignment Scores:
 Pred. No.: 6,86e-97 Length: 249
 Score: 1384.00 Matches: 247
 Percent Similarity: 99.60% Conservative: 1
 Best Local Similarity: 99.20% Mismatches: 1
 Query Match: 94.92% Indels: 0
 DB: 23 Gaps: 0
 US-09-598-982-20 (1-771) x AAU12011 (1-249)
 QY 7 CTCGAGAAAGAATGCTGCGGGGTCAGAGAGCGCCCGAGAGCAAGTGCCCTGCAGTGG 66
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 LEUGLULYSARGILEVALIGLYGLINGLUALAPROARGSERLYSTRPROTRPGLNVAL 20
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 67 AGCCTGAGATGCCAGGCGCCATACAGATGCACTTCTGCGGGGCTCCCTCATCCACCC 126
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 21 SERLEUARGVALHISGLYRPROTYRTRPMETHISPEHCYSGLYGLYSERLEULLEHISPRO 40
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 127 CAGTGGTCTCAGCCCGCGCGCTGCTGGGACCGGAGCTGAAGATCTGGCCGCTC 186
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 41 GLNTRPVALLEUTHRALALAHISCYVALDLPROMASPVALLYSBPVALALALEU 60
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 187 AGSGTGCACTGCGGGAGACAGACCTCTACACCGAGACACTGCTGCGGCTGCAGCAG 246
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 ARGVALGILNLEUARGLUSLHISLEUTHRYRGLINASPGLNLEUENUPROVALSERARG 80
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 247 ATCATCTGTGCACCAAGTTCTACACCGCCAGATCGAGCGAGCATCGCCCTGTGGAG 306
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 81 IIEILEVALHISPROGLINPHERYRTHRALAGLNLIEGLYALASPLLEALALEUENGLU 100
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 307 CTGGAGAGCCGCTGAAGTCTCCAGCCAGTCCACAGCGGTACACCTGCCCTGCTCA 366
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 101 LEUGLUGLUPROVALLYSPALISERSEHISVALHISTHVALTHRLEUPROVALSER 120
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 367 GAGACCTTCCCGCGGGGATGCCGTGCTGCTGCTGCTGCGGGGAGATGGAATGAT 426
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 GLUTHRPHETPROGLYMETPROCYSTRPVALTHNGLYTRPGLYASPVALASPSNAP 140
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 427 GAGCGCTCCACCGCATTTCTCTGAAGAGGTGAAGTCCCATATGAAAAACAC 486
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 141 GIUARGLEUPROPROBHEPROLEULYSGLINVALYVALPROLLEMETGLUASNHS 160
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 487 ATTTGTGACGCAAAATACCACTTGCGCGCTACAGGGAGAGAGAGCGCATCGTCCGT 546
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 161 IIECYSASPALALYSTYRTHISLEUGLYALATYRTHGLYASBPSPVALARGILEVALARG 180
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 547 GAGCAATCTGTGTGCCGGGAACACCGGAGGAGACTGTGCGAGGGCGACATCCGAGGG 606
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 ASPASPMELEUCYSALAGLYASNTHARGARGASPSETCYSGINGLYASPVALAGLYGLY 200
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 607 CCCCTGTGTGCAAGGTGAATGCACCTGGCTGCAGGCGGGCGGTGCTGAGTGGGCGAG 666
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 201 PROLEUVALCYSLYSVALASNGLYRTHRTPLEUGLINALAGLYVALVALSERTRPGLYGLU 220
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 667 GGCTGTGCCACCCCAACCGGCTGGCATCTACACCCGTGCTACCTACTGACTTGGACTGG 726
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 221 GLYCYSALAGLNPROMASNPTRPROGLYIETRYRTHARGVALTHRTYRTRLEUASPTRP 240
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 727 ATCCACCACTATGTCCCAAAAAAGCG 753
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 241 IIEHISHISTYRVALPROLYSLYSPRO 249
 CC |||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 6
 AAU12012
 ID AAU12012 standard; Protein: 249 AA.

[illegible]

Dd	41	GIINIRYALLETPIHIALAIAHISCYSVALIYPROSPAVALLYSASPIRELAIALALEU	60
Qy	187	AGGGGCAACTGCGGGAGACACACCTCTACACACGAGACACAGTGGCCGTCAGACAG	246
Dd	61	ARGVALIGINLEAARGIGINIHISLEUTYTRYGLIMSPGINLEULEUPROVALSERARG	80
Qy	247	ATCATGTGCACCCACACTTCTACACGGCCAGATGCGAGCGGACATGSCCTGGTGAG	306
Dd	81	ILEILEVALIHISPROGLINPHETRYTHRALAGINILEGLYLAASPILEALALEULEGLU	100
Qy	307	CTGAGAGAGCCGGTGAAGSTGTCCAGCACGTCACACAGGCAACCCCTGGCCCTCA	366
Dd	101	LEUGLUGIUPROVALYSVALISERSENHISVALHISTHYVALTHYLEUPRODALASER	120
Qy	367	GAGACSTTCCCCCGGGGATGCGCTGGTGCTACTGGTCGGGGGCGCATGTGGACAATGAT	426
Dd	121	GIUTHIRPHERPROFROGLYMERPROCYSTRYVALIHTHGLYTRPGIYASPVALAASPAASPR	140
Qy	427	GAGCGCCTCCACCGGCCATTTCTCTGAAGCAGGTGAAGTCCCATATGAAACACAC	486
Dd	141	GIUATGLEUPROFROPHROPHROLEULSGINVALLYSVALPROILEMETGLUASNHS	160
Qy	487	ATTGTGACGGAAATACACCTTGGGGCCCTACACGGGAGACAGCATGCCGATTCGTCCT	546
Dd	161	ILECYSAPRALALYSTYRHHSILEUGLYALATYRTHGLYASPARVALARGILEVALARG	180
Qy	547	GACGCAATGCTGTGGCGGGAGACACCGGAGGGGATCATGGCCAGGGCGCATCGGAGGG	606
Dd	181	ASPAHMETLEUCYSALAGLYASHTHARGARGASPERCYSGINGIYASPRALAGLYGLY	200
Qy	607	CCCTGTGTCTCAAGGTGAATGGCACCCTGGCTGCAGCGCGGCGTGTACACTGGGGCGAG	666
Dd	201	PROLEUVALCYLSYSVALASNGIYTHTRPREUGINLAAGLYVALISERTTPRGILU	220
Qy	667	GGCTGTGCCCAACCCACCGGCGCTGGCATCTACACCCGCTGACCTTACTTGGACTGG	726
Dd	221	GLYCYSALAGINPROASARGPROGLYILETYRTHARGVALTHTRYTYLEUASPTTP	240
Qy	727	ATCCACACTATGTGCCCAAAAGCGG 753	
Dd	241	ILEHSHISTRYVALPROLYLSYSPRO 249	
RESULT 7			
ID	AAU12006	standard; Protein: 249 AA.	
XX	AAU12006;		
AC	AAU12006;		
XX			
DT	09-APR-2002 (first entry)		
XX			
DE	Human beta-I tryptase.		
XX			
KW	Human; proteolytic tryptase; protease; beta-I tryptase; enzyme.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200198470-A2.		
XX			
PD	27-DEC-2001.		
XX			
PF	20-JUN-2001; 2001WO-US19681.		
XX			
PR	21-JUN-2000; 2000US-0598982.		
XX			
PA	(PROM-) PROMEGA CORP.		
XX			
PI	Maffit M, Niles AL, Haak-Frendscho M;		
XX	WPT; 2002-114578/15.		
DR	N-PSDB; AAS20763.		
XX			

PT DNA construct for producing enzymatically-inactive proteolytic
PT trypsinase, comprises DNA sequence encoding proteolytic trypsinase having
PT an active site mutation -

XX Disclosure: Page 74-75; 126pp; English.

CC The present invention relates to recombinant human proteolytic
CC trypsinases, active site mutants of these trypsinases and the methods for
CC producing these. The method involves the production of a DNA expression
CC construct comprising a promoter operably linked to a secretion signal
CC sequence which is operably linked to a DNA sequence encoding a
CC proteolytic trypsinase with an active site mutation (the construct drives
CC expression of a mature proteolytic trypsinase that lacks enzymatic activity
CC due to the active site mutation, in hosts transformed to contain the
CC construct). The method is useful for producing enzymatically-active
CC beta-II trypsinase. The active site mutants of proteolytic trypsinase provide
CC a tool to investigate the structural and functional properties of the
CC trypsinase and its enzymatic activity, and for modelling studies. The
CC enzymatically-active, recombinant proteolytic trypsinase produced are
CC useful as an antigen to generate anti-human trypsinase antibodies
CC and in drug screening for compounds which act as trypsinase inhibitors,
CC antagonists, agonists, etc. The present sequence represents human
CC beta-I trypsinase.

XX Sequence 249 AA:

Alignment Scores:

Prod. No.:	9,726-97	Length:	249
Score:	1382.00	Matches:	247
Percent Similarity:	99.20%	Conservative:	0
Best Local Similarity:	99.20%	Mismatches:	2
Query Match:	94.79%	Indels:	0
DB:	23	Gaps:	0

US-09-598-982-20 (1-771) x AAU12006 (1-249)

QY 7 CTCGAGAAAAAGATCGTCGGGGTCAGAGAGCCCGGAGAGCAAGTGGCCCTGGCAGTG 66
DB 1 LGGUULYARGIILEVALGILYGLNGUUAAPROVRSERYSTRPRORPGLNVAL 20
QY 67 AGCTGAGAGTCCAGGCGCCATCTGATGACACTTGGGGGGCTCCCTATCCACCC 126
DB 21 SerLeuAlrValHisclYProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
QY 127 CAGTGGGTCGTACCGCCCGGCGGTGCGTGGAGCGGAGTCAAGGATCGTGGCGCCCTC 186
DB 41 GINTPVALLeuThrAlaAlaHisCysValGlyProAspValIlyAspLeuAlaLeu 60
QY 187 AGGTCGCAACTGCGGAGCAGCAGCTCTACTACAGAGCAGCTGCTGGCGTCAGCAG 246
DB 61 ArgValGlnLeuAlrGlnGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
QY 247 ATCATGTCGACCCACAGTTTACACCGCCCGAGATCGAGCGGACATCGCCCTGTGAG 306
DB 81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGln 100
QY 307 CCGGAGAGCCGCGTCCAGGCTCCAGCGTCCACAGCGTCAACCTGGCCCTGCTCA 366
DB 101 LeuGlnGlnProValAsnValSerSerHisValHisThrValThrLeuProProAlaSer 120
QY 367 GAGACCTTCCCGCGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426
DB 121 GlnThrPheProProGlyMetProCysTrpValThrGlyTyrPheIlyAspValAspAsnAsp 140
QY 427 GAGCGCTCCACCGCCATTTCTCTGAAGCAGAGGTAGAGTCCCATTAATGAAAAACAC 486
DB 141 GlnAlrGlnProProProPheProLeuIlyGlnValIlyValProIleMetGlnAsnHis 160
QY 487 ATTGTGAGCAAAATATACACCTTGGCGCTACAGCGGAGAGAGAGTCCGCAATCGCCCT 546
DB 161 IleCysAspAlaIlySerTyrHisLeuGlyAlaTyrThrGlyAspAspAlaIlyIleValArg 180
QY 547 GAGACATGCTGTGTCGCGGAGAACCCGAGAGGACTCATGCGAGCGCATCCGAGCG 606

DB 181 AspAspMetLeuCysAlaGlnYasnThrArgArgAspSerCysGlnGlnIlySperGlyGly 200
QY 607 CCCCTGCTGTCGAAGTGAATGGACCTGCTGTCAGCGCGCGTGGTCACTGGGGCGAG 666
DB 201 ProLeuValCysIlySValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGln 220
QY 667 GGGTGGCCGAGCCGACCGGCTGGCATCTACACCGGTCAACCTACACTTGGACCTG 726
DB 221 GlnCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrIlyLeuAspTrp 240
QY 727 ATCCACCACTATGTCGCCCAAAAGCCG 753
DB 241 IleHisIleTyrValProIlySlyPro 249

RESULT 8

ID AAU12010 standard; Protein: 249 AA.

AC AAU12010;

DT 09-APR-2002 (first entry)

DE Human beta-II trypsinase active site mutant D91A #1.

KW Human: proteolytic trypsinase; protease; recombinant beta-II trypsinase;

KW enzyme; mutant; mutein.

OS Homo sapiens.

PN Synthetic.

PN WO200198470-A2.

PD 27-DEC-2001.

PF 20-JUN-2001; 2001WO-US19681.

PR 21-JUN-2000; 2000US-0598982.

PA (PROM-) PROMEGA CORP.

PI Maffit M, Niles AL, Haak-Frendscho M;

DR WPI: 2002-114578/15.

DR N-PSDB: AAS20776.

PT DNA construct for producing enzymatically-inactive proteolytic

PT trypsinase, comprises DNA sequence encoding proteolytic trypsinase having

PT an active site mutation -

PS Claim 8: Page 87-88; 126pp; English.

CC The present invention relates to recombinant human proteolytic
CC trypsinases, active site mutants of these trypsinases and the methods for
CC producing these. The method involves the production of a DNA expression
CC construct comprising a promoter operably linked to a secretion signal
CC sequence which is operably linked to a DNA sequence encoding a
CC proteolytic trypsinase with an active site mutation (the construct drives
CC expression of a mature proteolytic trypsinase that lacks enzymatic activity
CC due to the active site mutation, in hosts transformed to contain the
CC construct). The method is useful for producing enzymatically-active
CC beta-II trypsinase. The active site mutants of proteolytic trypsinase provide
CC a tool to investigate the structural and functional properties of the
CC trypsinase and its enzymatic activity, and for modelling studies. The
CC enzymatically-active, recombinant proteolytic trypsinase produced are
CC useful as an antigen to generate anti-human trypsinase antibodies
CC and in drug screening for compounds which act as trypsinase inhibitors,
CC antagonists, agonists, etc. AAU12009-AAU12024 represent recombinant
CC human beta-II trypsinase active site mutants.

XX Sequence 249 AA:

Alignment Scores:

Pred. No.: 1.64e-96 Length: 249
 Score: 1379.00 Matches: 247
 Percent Similarity: 99.20% Conservative: 0
 Best Local Similarity: 99.20% Mismatches: 2
 Query Match: 94.58% Indels: 0
 DB: 23 Gaps: 0

US-09-598-982-20 (1-771) x AAU12019 (1-249)

```

QY 7 CTCGAGAAAGATGTCGGGGGTGACAGAGGCCCCCAGAGCAATGGCCCTGACAGTG 66
DB 1 LeuGluYsArgIleValGIgYlGlnGlnAlaProAlaArgSerLysTrpProTrpIleVal 20
QY 67 AGCCTGAGAGTCCACGGCCCATCTGATGCATTCTGGGGGGCTCCCTCATCCAGCCC 126
DB 21 SerLeuArgValHisGIYProTYrTrpMetHisPheCysGIYSerLeuIleHisPro 40
QY 127 CAGTGGGTGTCAGCCCGCGCGGTGCGAGCCGAGCTGAAGATCTGGCCGCTC 186
DB 41 GlnTrpValIleThrAlaAlaHisCysValGIYProAspValLysAspLeuAlaIleu 60
QY 187 AGGTCGAACCTGCGGAGCAGCAGCTTACTACGAGCAGCAGCAGCTGCGGTCAGCAG 246
DB 61 ArgValGlnLeuArgGlnGlnHisLeuTYrTYrGlnAspGlnLeuLeuProValSerArg 80
QY 247 ATCATGTCGACCCACAGTTCTTACACCGCCCATGCGAGCGAGCATGCGCCCTGTGAG 306
DB 81 IleIleValHisProGlnPhenTYrThrAlaGlnIleGlyAlaAlaIleLeuLeuGln 100
QY 307 CTGGAGGAGCGGTTAGAGTCTCCAGCAGCAGCTCCACAGGAGTGCCTGCGCCCTCA 366
DB 101 LeuGlnGlnProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
QY 367 GAGACCTTCCCGCCGGGAGTCCGTGTCGTCGTCACCTGGCTGGGGGATGTGACAATGAT 426
DB 121 GlnTrpPheProProGlyMetProCysTrpValIleGlyTrpGlyAspValAspHis 140
QY 427 GAGCGCTCCACCGCCCATTTCTCTGAAGCAGGTGAAGTCCCATTAATGAAAAACAC 486
DB 141 GlnArgLeuProProPheProPheProLeuLysGlnValLysValProIleMetGlnHis 160
QY 487 ATTTGTGAGCAAAATACCACTTGGCGCTACAGCGGAGAGAGAGCTCGCATTCGTCCT 546
DB 161 IleCysAspAlaLysTrpHisLeuGlyAlaTYrThrGlyAspAspAlaArgIleValArg 180
QY 547 GACGACATCTGTGTCCGCGGAACAACCGGAGGACTCTATCCAGGCGACTCCGAGGCG 606
DB 181 AspAspMetLeuCysAlaGlnAsnThrArgAspSerCysGlnLysPserGlyGly 200
QY 607 CCCCTGTGTGCAAGGTGAATGCACTGCTGTCAGCGCGGCGGTGTCAGCTGGGGCGAG 666
DB 201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValLysSerTrpGlyGln 220
QY 667 GGCTGTGCCCCACCGCAACCGGCTGCACATCAACCGGTGACCACTGACTGTGACG 726
DB 221 GlnCysAlaGlnProAsnArgProGlyIleTYrThrArgValThrTYrTrpLeuAspTrp 240
QY 727 ATCCACACTATGTCCCAAAAAGCCG 753
DB 241 IleHisHisTYrValProLysLysPro 249

```

RESULT 9

AAU12019 ID AAU12019 standard; Protein: 249 AA.

AC AAU12019;

DT 09-APR-2002 (first entry)

XX Human beta-II tryptase active site mutant S194A #5.

XX Human; proteolytic tryptase; protease; recombinant beta-II tryptase;

KW enzyme; mutant; mutein.

```

XX XX Homo sapiens.
OS Synthetic.
XX XX WO200198470-A2.
PN 27-DEC-2001.
XX PD 20-JUN-2001; 2001WO-US19681.
XX PF 21-JUN-2000; 2000US-0598982.
XX PR (PROM-) PROMEGA CORP.
XX PA Maffit M, Niles AL, Haak-Frendscho M;
PI WPI: 2002-114578/15.
XX DR N-PSDB; AMS20785.
DR PT DNA construct for producing enzymatically-inactive proteolytic
PT tryptase, comprises DNA sequence encoding proteolytic tryptase having
PT an active site mutation -
XX PS Claim 8; Page 111-112; 126pp; English.
XX CC The present invention relates to recombinant human proteolytic
CC tryptases, active site mutants of these tryptases and the methods for
CC producing these. The method involves the production of a DNA expression
CC construct comprising a promoter operably linked to a secretion signal
CC sequence which is operably linked to a DNA sequence encoding a
CC proteolytic tryptase with an active site mutation (the construct drives
CC expression of a mature proteolytic tryptase that lacks enzymatic activity
CC due to the active site mutation, in hosts transformed to contain the
CC construct). The method is useful for producing enzymatically-active
CC beta-II tryptase. The active site mutants of proteolytic tryptase provide
CC a tool to investigate the structural and functional properties of the
CC protease and its enzymatic activity, and for modelling studies. The
CC enzymatically-active, recombinant proteolytic tryptase produced are
CC useful as an antigen to generate anti-human tryptase antibodies
CC and in drug screening for compounds which act as tryptase inhibitors,
CC antagonists, agonists, etc. AAU12009-AAU12024 represent recombinant
CC human beta-II tryptase active site mutants.
XX SQ Sequence 249 AA:

```

Alignment Scores:

Pred. No.: 1.64e-96 Length: 249
 Score: 1379.00 Matches: 246
 Percent Similarity: 99.20% Conservative: 1
 Best Local Similarity: 98.80% Mismatches: 2
 Query Match: 94.58% Indels: 0
 DB: 23 Gaps: 0

US-09-598-982-20 (1-771) x AAU12019 (1-249)

```

QY 7 CTCGAGAAAGATGTCGGGGGTGACAGAGGCCCCCAGAGCAATGGCCCTGACAGTG 66
DB 1 LeuGluYsArgIleValGIgYlGlnGlnAlaProAlaArgSerLysTrpProTrpIleVal 20
QY 67 AGCCTGAGAGTCCACGGCCCATCTGATGCATTCTGGGGGGCTCCCTCATCCAGCCC 126
DB 21 SerLeuArgValHisGIYProTYrTrpMetHisPheCysGIYSerLeuIleHisPro 40
QY 127 CAGTGGGTGTCAGCCCGCGCGGTGCGAGCCGAGCTGAAGATCTGGCCGCTC 186
DB 41 GlnTrpValIleThrAlaAlaHisCysValGIYProAspValLysAspLeuAlaIleu 60
QY 187 AGGTCGAACCTGCGGAGCAGCAGCTTACTACGAGCAGCAGCTGCGGTCAGCAG 246
DB 61 ArgValGlnLeuArgGlnGlnHisLeuTYrTYrGlnAspGlnLeuLeuProValSerArg 80
QY 247 ATCATGTCGACCCACAGTTCTTACACCGCCCATGCGAGCGAGCATGCGCCCTGTGAG 306

```


|||||
 Db 221 GlycylalaglnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAsprrp 240
 Oy 727 ATCCACGACTATGTCGCCCAAAAGCG 753
 |||||
 Db 241 IIEHISHSTYrValProLysLysPro 249

RESULT 11
 AAU12013
 ID AAU12013 standard: Protein: 245 AA.

AC AAU12013:

DT 09-APR-2002 (first entry)

DE Human beta-II tryptase active site mutant H44A #2.

KM Human; proteolytic tryptase: protease; recombinant beta-II tryptase;
 KW enzyme; mutant; muteln.

OS Homo sapiens.

OS Synthetic.

PN WO200198470-A2.

XX 27-DEC-2001.

PF 20-JUN-2001; 2001WO-US19681.

PR 21-JUN-2000; 2000US-0598982.

PA (PROM-) PROMEGA CORP.

PI Maffit M, Niles AL, Haak-Frendscho M;

DR WPI: 2002-114578/15.

DR N-PSDB: AAS20779.

PT DNA construct for producing enzymatically-inactive proteolytic
 PT tryptase, comprises DNA sequence encoding proteolytic tryptase having
 PT an active site mutation -

XX Claim 40; Page 95-96; 126pp; English.

XX The present invention relates to recombinant human proteolytic
 CC tryptases, active site mutants of these tryptases and the methods for
 CC producing these. The method involves the production of a DNA expression
 CC construct comprising a promoter operably linked to a secretion signal
 CC sequence which is operably linked to a DNA sequence encoding a
 CC proteolytic tryptase with an active site mutation (the construct drives
 CC expression of a mature proteolytic tryptase that lacks enzymatic activity
 CC due to the active site mutation, in hosts transformed to contain the
 CC construct). The method is useful for producing enzymatically-inactive
 CC beta-II tryptase. The active site mutants of proteolytic tryptase provide
 CC a tool to investigate the structural and functional properties of the
 CC protease and its enzymatic activity, and for modeling studies. The
 CC enzymatically-active, recombinant proteolytic tryptase produced are
 CC useful as an antigen to generate anti-human tryptase antibodies
 CC and in drug screening for compounds which act as tryptase inhibitors,
 CC antagonists, agonists, etc. AAU12009-AAU12024 represent recombinant
 CC human beta-II tryptase active site mutants.

XX Sequence 245 AA;

Alignment Scores:

Pred. No.: 3.92e-96 Length: 245
 Score: 1374.00 Matches: 245
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 94.24% Indels: 0
 DB: 23 Gaps: 0

US-09-598-982-20 (1-771) x AAU12013 (1-245)

Oy 19 ATCTGTGGGGGTCCAGAGGCCCCAGAGACAGTGGCCCTGGAGCTGAGAGTC 78
 |||||
 Db 1 IIEVAlGlgYlGlnGlnAlaIarProArgSerIySTrProTrpGlnValSerIeuaRgVal 20
 Oy 79 CAGGCCCACTAGTGAATGCATCTTGGCGGGGCTCCCTCATCCACCCCACTGGGTGCTG 138
 |||||
 Db 21 HISGlyProTyrTrpMetHisPheCysGlySerIeIleHisProGlnTrpValIleu 40
 Oy 139 ACCGGCGGGGCTGGTGGAGCAGCTCAGAGATCTCGCCGCTCAGAGGTGACATG 198
 |||||
 Db 41 ThrAlaAlaAlaCysValGlyProAspValLysAspIleuAlaIaIeuaRgValGlnIleu 60
 Oy 199 CCGAGACAGACACTCTACTACAGAGACAGATGCTGCGCGGTGACAGATCATCTGCAC 258
 |||||
 Db 61 ArgGlnGlnHisIeulTyrTyrGlnAspGlnIleuLeuProValSerArgIleIleValHis 80
 Oy 235 CCACAGTTCTACACGCGCCACATCGAGCGGACATCGCCCTGCTGGAGCTGAGAGCGCG 318
 |||||
 Db 81 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaIeulIeulIeulIeulPro 100
 Oy 319 GTGAAGGTCTCCAGCCAGCTCCACAGCGGTACCGCCCTCGCTCAGAGACTTTCGCC 378
 |||||
 Db 101 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGlnThrPhePro 120
 Oy 379 CCGGGATGCGCTGCTGGGTCACTGCTGGGCGAGTGTGACAAATGATGAGCGCTCCCA 438
 |||||
 Db 121 ProGlyMetProCysTrpValThrGlnTrpGlyAspValAspAsnAspIleuArgLeuPro 140
 Oy 439 CCGCATTTCTCTGTAGACAGGTGAAGTCCCATATGAAACACACATTTGTGACGCA 498
 |||||
 Db 141 ProporphProLeuIleuLysGlnValLysValProIleMetGlnHisIleCysAspAla 160
 Oy 499 AATATCCACCTGGGGCTGACACAGGAGACAGCGATGCTGCTGACACATGCTGCTG 558
 |||||
 Db 161 LysTyrHisLeuGlyAlaIaIaTyrThrLysAspValArgIleValAlaIaIaIaIaIaIa 180
 Oy 559 TGTGCGGGGAGACACCCGAGGAGCTCATGCCAGGGCGACTCCGAGGGCCCTGGTGTGC 618
 |||||
 Db 181 CysAlaGlyAsnThrThrArgArgAspSerCysGlnGlyAspSerCylGlyProIeuaLys 200
 Oy 619 AAGTGAATGACACTGCTGCTGAGCGGGGCTGTGACGTGGGCGAGAGGCTGTGCCAG 678
 |||||
 Db 201 LysValAsnGlyThrTrpLeuGlnAlaGlyValIaIaIaIaIaIaIaIaIaIaIaIa 220
 Oy 679 CCCAACCGGCGTGGATCTACACCGGTGCACCTACTCTTGAGACTGGATCCACACTAT 738
 |||||
 Db 221 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisIlyr 240
 Oy 739 GTCCCAAAAAGCCG 753
 |||||
 Db 241 ValProLysLysPro 245

RESULT 12
 AAU12018
 ID AAU12018 standard: Protein: 249 AA.

XX AAU12018:
 AC AAU12018:
 DT 09-APR-2002 (first entry)
 DE Human beta-II tryptase active site mutant D91A #3.
 DE Human; proteolytic tryptase: protease; recombinant beta-II tryptase;
 KW enzyme; mutant; muteln.
 OS Homo sapiens.
 OS Synthetic.
 PN WO200198470-A2.
 XX 27-DEC-2001.
 PD

PF 20-JUN-2001; 2001WO-US19681.
 XX
 PR 21-JUN-2000; 2000US-0598982.
 XX
 PA (PROM-) PROMEGA CORP.
 XX
 PI Maffit M, Niles AL, Haak-Frendscho M;
 XX
 DR WPI: 2002-114578/15.
 DR N-PSDB: AAS20784.
 XX
 PT DNA construct for producing enzymatically-inactive proteolytic
 PT trypsinase, comprises DNA sequence encoding proteolytic trypsinase having
 PT an active site mutation -
 XX
 PS Claim 8; Page 109-110; 126pp; English.
 CC
 CC The present invention relates to recombinant human proteolytic
 CC trypsinases, active site mutants of these trypsinases and the methods for
 CC producing these. The method involves the production of a DNA expression
 CC construct comprising a promoter operably linked to a secretion signal
 CC sequence which is operably linked to a DNA sequence encoding a
 CC proteolytic trypsinase with an active site mutation (the construct drives
 CC expression of a mature proteolytic trypsinase that lacks enzymatic activity
 CC due to the active site mutation, in hosts transformed to contain the
 CC construct). The method is useful for producing enzymatically-active
 CC beta-II trypsinase. The active site mutants of proteolytic trypsinase provide
 CC a tool to investigate the structural and functional properties of the
 CC protease and its enzymatic activity, and for modelling studies. The
 CC enzymatically-active, recombinant proteolytic trypsinase produced are
 CC useful as an antigen to generate anti-human trypsinase antibodies
 CC and in drug screening for compounds which act as trypsinase inhibitors,
 CC antagonists, agonists, etc. AAU12009-AAU12024 represent recombinant
 CC human beta-II trypsinase active site mutants.
 XX
 SQ Sequence 249 AA:

Alignment Scores:
 Pred. No.: 3,93e-96 Length: 249
 Score: 1374.00 Matches: 246
 Percent Similarity: 98.80% Conservative: 0
 Best Local Similarity: 98.80% Mismatches: 3
 Query Match: 94.24% Indels: 0
 DB: 23 Gaps: 0

US-09-598-982-20 (1-771) x AAU12018 (1-249)

QY 7 CTCGAGAAAAGATGCTCGGGGTCAGAGAGCCCGCAGAGCAAGTGGCTTGCAGTG 66
 DB 1 LGGGulysArgIleValIcylGlnGlnAlaProArgSerLysTrpProIrrpGlnVal 20
 QY 67 AGCTGAGAGTCCAGCGCCCATCTGATGATGACCTTCTGGGGGCTCCCTCATCCAGCC 126
 DB 21 SerLeuArgValIhISglYProTyrTrpMetChISphecYsGlyGlySerLeuIleHisPro 40
 QY 127 GAGTGGGCTGACCGCGCGCGCTGCGTGGAGCCGAGAGTCAAGATGTGCGCGCCCTC 186
 DB 41 GlnTrpValLeuThrAlaIhIScysValIcylProAspValLysPseuAlaIleu 60
 QY 187 AGGTTGCAACTCGGGAGAGCAGCCTCTACTACAGAGCAGCAGCTGCTGGCTGACAGG 246
 DB 61 ArgValGlnLeuArgLcylGlnIleHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
 QY 247 ATCATGCTGACCCACAGTTCTTACACCGCCCAAGTGGAGCGGACATCGCTGTGAG 306
 DB 81 IleIleValIhISProGlnPheTyrThrIleGlnIleGlyAlaIleAlaIleLeuLeuGln 100
 QY 307 CCGGAGGACCGGTGAAGTCTCCAGCCAGTCCACAGCGTACACCTGTGCGCTGCTCA 366
 DB 101 LeuGlnGlnProValAsnValSerSerHisValIhISThrValIhISLeuProValSer 120
 QY 367 GAGACCTTCCCGCGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 121 GlnThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
 QY 427 GAGCGCTCCCGCCGCGCATTTCTCTGAAAGCAGGTGAAGTCCCATATATGAAAAACAC 486
 DB 141 GlnArgLeuProProPheProPheProLeuLysGlnValLysValProIleMetIuAsnHis 160
 QY 487 ATTTGTGAGCAAAATACACCTTGCGGCTACAGGAGAGACAGTCCGATCGTCCGT 546
 DB 161 IleCysAspAlaIleLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
 QY 547 GAGCAGATGCTGTGTCCGCGGAACACCGGAGGAGCTATCCAGGCGCATCTCCGAGGG 606
 DB 181 AspAspMetLeuCysAlaIleLysAsnThrArgAspSerCysGlnGlyAspSerGlyGly 200
 QY 607 CCGCTGGTGTGCAAGGTGATGGCACCTGGCTGACAGCGCGGCTGTGCTGAGTGGGCGAG 666
 DB 201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGln 220
 QY 667 GCGTGGCCGAGCCGCGGCTGGCATCTACACCGCGTGCACCTACTTGGACGTG 726
 DB 221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTrpLeuAspTrp 240
 QY 727 ATCCACCACTATGTCCCAAAAAGCG 753
 DB 241 ILeHISHTYrValProLysLysPro 249
 RESULT 13
 AAU12021
 ID AAU12021 standard; Protein: 245 AA.
 XX
 XX AAU12021;
 AC
 XX 09-APR-2002 (first entry)
 DT
 XX
 DE Human beta-II trypsinase active site mutant H44A #4.
 DE
 XX Human: proteolytic trypsinase: protease: recombinant beta-II trypsinase:
 KW enzyme; mutant; mutein.
 KW
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200198470-A2.
 XX
 PD 27-DEC-2001.
 PD
 PF 20-JUN-2001; 2001WO-US19681.
 XX
 PR 21-JUN-2000; 2000US-0598982.
 XX
 PA (PROM-) PROMEGA CORP.
 XX
 PI Maffit M, Niles AL, Haak-Frendscho M;
 XX
 DR WPI: 2002-114578/15.
 DR N-PSDB: AAS20787.
 XX
 PT DNA construct for producing enzymatically-inactive proteolytic
 PT trypsinase, comprises DNA sequence encoding proteolytic trypsinase having
 PT an active site mutation -
 XX
 PS Claim 40; Page 117-118; 126pp; English.
 CC
 CC The present invention relates to recombinant human proteolytic
 CC trypsinases, active site mutants of these trypsinases and the methods for
 CC producing these. The method involves the production of a DNA expression
 CC construct comprising a promoter operably linked to a secretion signal
 CC sequence which is operably linked to a DNA sequence encoding a
 CC proteolytic trypsinase with an active site mutation (the construct drives
 CC expression of a mature proteolytic trypsinase that lacks enzymatic activity
 CC due to the active site mutation, in hosts transformed to contain the
 CC construct). The method is useful for producing enzymatically-active
 CC beta-II trypsinase. The active site mutants of proteolytic trypsinase provide


```

QY 139 ACCGCGCGGCGTGGTGAGCGGACGTCAAGATCTGGCGCCCTCAGGGTGCACATG 198
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 41 ThrAlaAlaHisCysValGlyProAspValLysAspLeuAlaAlaLeuArgLysGlnLeu 60
QY 199 CGGAGACACACCTTACTACAGAGACAGCTGCTGCCGCTCAGACAGATCATCTGCAC 258
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 61 ArgGlnGlnHisLeuTyrTrpGlnAspGlnLeuProValSerArgLysIleValHis 80
QY 259 CCACAGTTCTACACCGCCAGATCCGAGCGGACATCCGCCCTCTGGAGCTGGAGAGCGG 318
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 81 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlnLeuGlnPro 100
QY 319 GTGAAGGTCTCAGCAGCATCCACAGGTCACCTGCCCTGCCCTCAGAGACCTTCCCC 378
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 101 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGlnThrPhePro 120
QY 379 CCGGGGATCCCGTGTGGTCACTGGCTGGGGCGATGTGGACATGATGAGCGCCCTCCA 438
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 121 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGlnArgLeuPro 140
QY 439 CGGCCATTTCCCTGAGACAGGTGAAGGTCCCATATGGAACACACATTTGTGACGA 498
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 141 ProProPheProLeuLysGlnValLysValProIleMetGlnAsnHisIleCysAspAla 160
QY 499 AATATCACCTTGGCGCTACACAGCGGAGACAGATCCGCATCGCTCCGTGACGACATGCTG 558
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 161 LysTyrHisLeuGlyAlaLysThrGlyAspAspValArgIleValArgAspAspMetLeu 180
QY 559 TGTGCGGGAGACACCGGAGGAGCATCATCCAGGCGACCTCCGAGAGGCCCTGTGTGTC 618
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 181 CysAlaGlyAsnThrArgTrpAspSerCysGlnGlyAspSerGlyLysProLeuValCys 200
QY 619 AAGGTGAATGGACCTGGCTGCAGCGGGCGCTGGTACGCTGGGGCGAGGGCTGTGCCCA 678
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 201 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpLysGlnCysAlaGln 220
QY 679 CCCAACCGGCTGGCATCTACACCGCTGCACCTACTACTGAGCTGATCCACCACTAT 738
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 221 ProAsnArgProGlyLysTrpLeuThrArgValThrTyrTyrLeuAspTrpIleHisLysTyr 240
QY 739 GTCCCAAAAAGCCG 753
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 241 ValProLysLysPro 245

RESULT 15
AAM64240
ID AAM64240 standard; Protein; 274 AA.
XX
AC AAM64240;
XX
DT 24-NOV-1998 (first entry)
XX
DE Human mast cell tryptase II/beta.
XX
KW Mast cell tryptase II/beta; human; MCP-7; mast cell protease 7;
KW blood clot; anticoagulant; myocardial infarction; reocclusion;
KW thromboembolism; cerebral embolism; thrombosis; therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..18
FT /label= Sig_peptide
FT Peptide 19..28
FT /label= Pro_peptide
FT Protein 29..274
FT /label= Mat_protein
XX
PN MO9824886-A1.
XX
PD 11-JUN-1998.
XX

```

```

PF 25-NOV-1997; 97WO-US21620.
XX
PR 04-DEC-1996; 96US-0032354.
XX
PA (BGHM) BRIGHAM & WOMEN'S HOSPITAL.
XX
PI Stevens RL.
XX
DR WPI: 1998-33308/29.
DR N-PSDB: AAV44330.
XX
PT New compositions containing tryptase-7, e.g. mouse mast cell
PT protease-7 - are used to treat clot formation in e.g. myocardial
PT infarction, reocclusion following angioplasty or pulmonary
PT thromboembolism
XX
PS Disclosure: Page 66-67; 92pp; English.
XX
CC This is the deduced amino acid sequence of human mast cell tryptase
CC II/beta (see also AAV44330). The invention provides: compositions
CC comprising an isolated tryptase-7 that may include chimeric proteins
CC that contain (a) a human tryptase for all but the active site region
CC and (b) the substrate-binding pocket of mouse tryptase-7 or its
CC homologues (see AAM64233-39); a method for treating a blood clot by
CC administering a nucleic acid molecule that codes for a tryptase-7,
CC or an expression product, to decrease fibrinogen activity; a nucleic
CC acid encoding a serine protease (SP) and a method of producing a
CC mature SP by expressing the inactive zymogen in a host cell, and
CC cleaving the enterokinase susceptibility domain. The tryptase-7
CC polypeptides can be used to treat disorders mediated by undesirable
CC thrombus clot formation such as myocardial infarction and
CC reocclusion following angioplasty of blood clots associated with
CC pulmonary thromboembolism, deep vein thrombosis, cerebral embolism,
CC renal vein and peripheral arterial thrombosis. They are also
CC useful for all surgical procedures that require decreased blood
CC clots.
XX
SQ Sequence 274 AA:

Alignment Scores:
Pred. No.: 1,13e-95 Length: 274
Score: 1368.00 Matches: 244
Percent Similarity: 99.59% Conservative: 0
Best Local Similarity: 99.59% Mismatches: 1
Query Match: 93.83% Indels: 0
DB: 19 Gaps: 0

US-09-598-982-20 (1-771) x AAM64240 (1-274)
QY 19 ATGCTGGGGGTGTCAGAGAGCGCCCGACAGACAGTGGCGAGGTGAGAGATC 78
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 30 IleValGlyGlyGlnAlaProArgSerLysTrpProTrpGlnValSerLeuAlaGlyAl 49
QY 79 CACGGCCCATACTGATGACATTCTGCGGGGGCTCCCTCATCCACCCAGTGGGTGCTG 138
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 50 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 69
QY 139 ACCGCGCGGCGTGGCTGGGACCGGACGTCAAGATCTGGCGCCCTCAGAGGTGCACAG 198
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 70 ThrAlaAlaHisCysValGlyProAspValLysAspLeuAlaAlaLeuArgLysGlnLeu 89
QY 199 CGGAGACACACCTTACTACAGAGACAGCTGCTGCCGCTCAGAGATCATCTGCAC 258
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 90 ArgGlnGlnHisLeuTyrTrpGlnAspGlnLeuProValSerArgLysIleValHis 109
QY 259 CCACAGTTCTACACCGCCAGATCCGAGCGGACATCCGCCCTCTGGAGCTGGAGAGCGG 318
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 110 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlnLeuGlnPro 129
QY 319 GTGAAGGTCTCAGCAGCATCCACAGGTCACCTGCCCTGCCCTCAGAGACCTTCCCC 378
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 130 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGlnThrPhePro 149

```

```
QY 379 CCGGGATGCGCTGCTGGGTCACTGGCTGGGGGATGTGACATGATGAGCGCCYCCCA 438
    |||||||
Db 150 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 169
    |||||||
QY 439 CCGCCATTTCCTCTGTAAGCAGGTGAAGTCCCATATATGGAANAACCATTTGTGACGA 498
    |||||||
Db 170 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 189
    |||||||
QY 499 AATATCCACCTTGGCGCCCTACACGGGAGACAGTCCGATGCTCCGTGACGACATGCTG 558
    |||||||
Db 190 LysTyrHisLeuGlyAlaIleThrGlyAspAspValArgIleValArgAspAspMetLeu 209
    |||||||
QY 559 TGTGCCGGGAGACACCCGAGGAGACTCATGACAGGGGACTCCGAGGGCCCTGGTGTGC 618
    |||||||
Db 210 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 229
    |||||||
QY 619 AAGGTGATGGCACCCTGGCTGCAGGGCGGCTGTGTCAGCTGGGGCGAGGGCTGTGCCAG 678
    |||||||
Db 230 LysValAsnGlyThrTrpLeuGlnAlaGlyValIleSerTrpGlyGlnGlyCysAlaGln 249
    |||||||
QY 679 CCCAACCGGCTGGCATCTACACCCGCTGTACCTACTACTGTGAGCTGATCCACCATCTAT 738
    |||||||
Db 250 ProAsnArgProGlyIleTyrThrArgValIleTyrTrpLeuAspTrpIleHisIleTyr 269
    |||||||
QY 739 GTCCCCAAAAAGCCG 753
    |||||||
Db 270 ValProLysLysPro 274
```

Search completed: January 31, 2003, 06:53:08
Job time : 59 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: January 31, 2003, 06:54:15 : Search time 31 seconds
(without alignments)
1003.720 Million cell updates/sec

Title: US-09-598-982-20

Perfect score: 1458
Sequence: 1 gggcccccgcgagaaagaat.....cgtgaagcgcgcgcgcgtcgt 771

Scoring table:

BLOSUM62			
Xgapop 10.0 ,	Xgapext 0.5		
Ygapop 10.0 ,	Ygapext 0.5		
Fgapop 6.0 ,	Fgapext 7.0		
Delop 6.0 ,	Delext 7.0		

Searched: 122226 segs, 20178551 residues

Total number of hits satisfying chosen parameters: 244452

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing:

Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US05958982/runat_27012003_073718_19634/app_query.fasta.1.967
-DB=Published_Applications_AA -OFM=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODB=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US05958982@cgn1.1.3_@runat_27012003_073718_19634
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA:*

1:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	608.5	41.7	311	10	US-09-900-754-2
2	569.5	39.1	288	10	US-09-888-615-115
3	563.5	38.6	290	9	US-10-041-006A-7
4	563.5	38.6	290	9	US-10-028-072-222

5	563.5	38.6	290	12	US-10-040-655-7	Sequence 7, Appli
6	554.5	38.0	315	9	US-10-041-006A-9	Sequence 9, Appli
7	554.5	38.0	315	12	US-10-040-655-9	Sequence 9, Appli
8	536.5	36.8	284	12	US-10-041-400A-7	Sequence 7, Appli
9	536.5	36.8	284	12	US-10-041-264A-7	Sequence 7, Appli
10	536.5	36.8	284	12	US-10-042-091A-7	Sequence 7, Appli
11	527.5	36.2	316	12	US-10-041-400A-9	Sequence 9, Appli
12	527.5	36.2	316	12	US-10-041-264A-9	Sequence 9, Appli
13	527.5	36.2	316	12	US-10-042-091A-9	Sequence 9, Appli
14	512	35.1	343	10	US-09-948-094-2	Sequence 2, Appli
15	504.5	34.6	238	9	US-10-067-761-19	Sequence 19, Appli
16	504.5	34.6	238	10	US-09-804-156-19	Sequence 19, Appli
17	501.5	34.4	317	9	US-09-905-291A-263	Sequence 263, App
18	501.5	34.4	317	9	US-10-040-803-7	Sequence 7, Appli
19	501.5	34.4	317	9	US-09-902-853-263	Sequence 263, App
20	501.5	34.4	317	9	US-09-907-824-263	Sequence 263, App
21	501.5	34.4	317	9	US-09-907-841-263	Sequence 263, App
22	501.5	34.4	317	9	US-09-904-011-263	Sequence 263, App
23	501.5	34.4	317	10	US-09-909-320-263	Sequence 263, App
24	501.5	34.4	317	10	US-09-909-088B-263	Sequence 263, App
25	501.5	34.4	325	10	US-09-908-711-114	Sequence 114, App
26	501.5	34.4	325	10	US-09-764-898-278	Sequence 278, App
27	501.5	34.4	351	10	US-09-764-898-207	Sequence 207, App
28	499.5	34.3	322	10	US-09-764-898-279	Sequence 279, App
29	494	33.9	314	9	US-09-905-291A-257	Sequence 257, App
30	494	33.9	314	9	US-09-902-853-257	Sequence 257, App
31	494	33.9	314	9	US-09-907-824-257	Sequence 257, App
32	494	33.9	314	9	US-09-907-841-257	Sequence 257, App
33	494	33.9	314	9	US-09-904-011-257	Sequence 257, App
34	494	33.9	314	10	US-09-909-320-257	Sequence 257, App
35	494	33.9	314	10	US-09-885-441-14	Sequence 14, Appli
36	494	33.9	314	10	US-09-909-088B-257	Sequence 257, App
37	492.5	33.8	327	9	US-10-040-803-8	Sequence 8, Appli
38	477.5	32.8	818	10	US-09-888-615-111	Sequence 111, App
39	476	32.6	453	9	US-09-978-197A-69	Sequence 69, Appli
40	476	32.6	453	9	US-09-978-169-69	Sequence 69, Appli
41	476	32.6	453	9	US-09-978-197A-69	Sequence 69, Appli
42	476	32.6	453	9	US-09-999-832A-69	Sequence 69, Appli
43	476	32.6	453	9	US-09-978-189-69	Sequence 69, Appli
44	476	32.6	453	9	US-10-174-590-64	Sequence 64, Appli
45	476	32.6	453	9	US-10-176-758-64	Sequence 64, Appli

ALIGNMENTS

RESULT 1
US-09-900-754-2
; Sequence 2, Application US/09900754
; Patent No. US20020026654A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Keith D.
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING TRYPTASE GENE
; FILE REFERENCE: R-372
; CURRENT APPLICATION NUMBER: US/09/900,754
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US 60/216,109
; PRIOR FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: US 60/223,172
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 60/244,111
; PRIOR FILING DATE: 2000-10-26
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-900-754-2
Alignment Scores: 5.56e-33 Length: 311
Pred. No.: 311

Score: 608.50 Matches: 123
Percent Similarity: 60.63% Conservative: 31
Best Local Similarity: 48.43% Mismatches: 81
Query Match: 41.74% Indels: 19
DB: 10 Gaps: 6

US-09-598-982-20 (1-771) x US-09-900-754-2 (1-311)

```
OY 16 AGATCGTCGCGGGGCTCAGAGCGCCCGAGAGCAAGTGGCTGCGAGGTGAGCTGAGA 75
DB 29 AGTlEValGlglylHlshlAlAlAlProAlaglyThrPrProTrpGlAlAlAserleuArg 48
OY 76 GTCCAGCGGCCCTACTGATGCACCTTCTGCGGGGCTCCCTCATCCACCCAGTGGGTG 135
DB 49 leuHlslYs-----ValHlslValCysglYglYserleuLeuSerProcllunTrpVal 65
OY 136 CTGACCGCGCGCGCTGGCTGGAGCCGAGCTCAAGATCTGCGCGCCCTCAGGGGTGCAA 195
DB 66 leuThrlAlAlAlHlshlscYsheserGlYserValAsnSer---SerAspTrpGlAlValHlsh 84
OY 196 CTGCGGGAGCAGCAGCTCTACTACAGACAGCAGTGTGCTCCGCTCAGCAGGATCATGCTG 255
DB 85 leuGlYglYleuThrlValThrlleuSerProHlshPheSerThrlValYslArglllelleMet 104
OY 256 CACCCACAGTTCTACACCGCCAGATCGGA-----CGGCACATCGCCCTG 300
DB 105 -----TyrThrlglYserProglYrProProglYserSerGlYslAspIleAlleu 120
OY 301 CTGAGAGCTGGAGAGCCGCTGAAAGTCTCCAGCCAGCTCCACAGGTCACCTGCCCTC 360
DB 121 ValGlnleuSerSerProValAlAlleuSerSerGlAlValGlnProValCysleuProGlu 140
OY 361 GCGTCAGAGACCTTCCCCCGGGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 141 AlAserAlAlAsPheThrlYrProglYmetGlnCysTrpValThrlclYTrpGlYThrGly 160
OY 421 AATGATGAGCGGCTCCACCGCCATTTCTCTGAAGCAGGTAGAGTCCCATATGAA 480
DB 161 GluGlYglYrProleuYrProProTrpAsnleuGlnAlAlValserValAlAsp 180
OY 481 AACCACTTTGTGACGGCAAAATACCACTTGGCGCTTACCGGGAGACAGCAGCTCCCATC 540
DB 181 ValYsThrlCysSerGlN-----AlatYrAsnSerProAsnGlYserleu 195
OY 541 GTCCGTGACGATGCTGTGTGCGGAGACACCCGAGGAGCTCATGCGGCGGCTGCC 600
DB 196 lIeGlnPrProAspMetleuCysAlArGlyPro---GlyAsPAlAcYsGlnAspAspSer 214
OY 601 GGAGGCGCCCTGTGTGCAAGGTGAATGGCACCTGGCTGCGAGCGGGCGTGTCACTGG 660
DB 215 GlYglYrProleuValCysGlnValAlaglyThrTrpGlnGlnAlAlglYAlValAserTrp 234
OY 661 GCGGAGGCTGTGGCCAGCCCAACCGGCTGGCATCTACACCGGTCTACCTACTG 720
DB 235 GlYglYglYcYsGlyArGrProAspArGrProGlYValYThrlAlArGrValThrlAlArYVal 254
OY 721 GACTGATCCACCACTATCCCAAAAGCCGTAGAGCGG 762
DB 255 AsnTrpIleHlshlshlslleProGlAlAlaglyglYserGlY 268
```

RESULT 2
US-09-888-615-115
Sequence 115, Application us/09888615
Patent No. US20020064856A1

GENERAL INFORMATION:
APPLICANT: PLOWMAN, GREGORY
APPLICANT: WHYTE, DAVID
APPLICANT: CAENEPEEL, SEAN
APPLICANT: CHARYDCZAK, GLEN
APPLICANT: MANNING, GERRARD
APPLICANT: SUDARSANAM, SUCHA
TITLE OF INVENTION: NOVEL PROTEASES
FILE REFERENCE: 038602/1214

CURRENT APPLICATION NUMBER: US/09/888, 615
CURRENT FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 60/214, 047
PRIOR FILING DATE: 2000-06-26
NUMBER OF SEQ ID NOS: 150
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 115
LENGTH: 288
TYPE: PRF
ORGANISM: Homo sapiens
US-09-888-615-115

Alignment Scores:

Pred. No.: 1,98e-30 Length: 288
Score: 569.50 Matches: 116
Percent Similarity: 58.23% Conservative: 29
Best Local Similarity: 46.59% Mismatches: 67
Query Match: 39.06% Indels: 37
DB: 10 Gaps: 5

US-09-598-982-20 (1-771) x US-09-888-615-115 (1-288)

```
OY 19 ATCGTCGCGGGGCTCAGAGAGCCCGAGAGCAAGTGGCCCTGCGAGGTGAGAGTGC 78
DB 73 lIeValGlglylHlshlAsnAlArProglYsTrPrProTrpGlAlValserleuArYVal 92
OY 79 -----CACGGCCCATACTGATGACACTTCTGCGGGGCTCCCTCATCCACCCCGAG 129
DB 93 TyrSerThrlHlshlTrAlAserTrAlAlHlshlIleCYsGlglylYserleuHlshlProGln 112
OY 130 TGGGTGCTGACCGCGCGCGGCTGGCTGGAGCCGAGCTCAAGATCTGGCGCCCTCAG 189
DB 113 TrpValleuThrlAlAlAlHlshlscYslIlePheTrpYsAspThrlAspProserlIeYrArg 132
OY 190 CTGCACACTCGCGGAGCAGCAGCAGCTACTACAGAGACAGCAGCAGTCTGCGGTACAGAGATC 249
DB 133 lIeHlslAlAglyAspAlYrleuYrGlylYrArgGlyleuLeuAsnValserHrGlie 152
OY 250 ATCGTCGACCCACAGTCTTCTACACCGCCAGATCGAGCGGAGCATCGCCCTGCTGAGCTG 309
DB 153 lIeValHlshlProAsnTrYrAlThrlArGlYleuGlyAlAlAspValAlAlleuLeuGlnleu 172
OY 310 GAGGAGCCGCTGAAGGTCTCCAGCCAGCTCACACAGGTCAACCTGCCCTGCCCTCAGAG 369
DB 172 ----- 172
OY 370 ACCTTCCCGCCCGGGATGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 429
DB 173 -----ProglYserProleu-----serProGlu 180
OY 430 CCGTCCCGCACCGCATTTCTCTGAAGCAGGTGAAGTCCCATATGGAACCAACAT 489
DB 181 SerleuProProProTrpArgleuGlnGlnAlAlAserValGlnValleuGlnAsnAlAlVal 200
OY 490 TGTGACGCAAAATAC---CACCTTGGCGCTTACCGGAGAGACAGTCCGATGCTCGT 546
DB 201 CysGlYglYrProTrYrArgAsnAlAserGlYHlshlGlyAspArg---GlnleuIleleu 219
OY 547 GACGAGATCCTGTGTGCGCGGAAACACCCGAGGAGACTATGCCAGGCGGACATCCGAGGG 606
DB 220 AspAspMetleuCysAlAlYserGlYglYrArgAspSerCysTrpGlYslYsPserGlYglY 239
OY 607 CCCCTGTGTGCAAGGTGAATGCACTGTGCTGCGAGCGGCGGTGCTAGCTGAGCGGAG 666
DB 240 ProleuValCysArgleuArGlglyserTrpArgleuValAlGlyAlValserTrpGlYTr 259
OY 667 GCGTGTGCCACCGCAACCGGCTGGCATCTACACCGGTGTCACTTACTTGGACTGG 726
DB 260 GlYcYsThrlleuArGAsrPheProglYValYThrlHlshlValGlnlIeYrValleuTrp 279
OY 727 ATCCACACTATGTCCCAAAAAGCCG 753
DB 280 lIeIeuglnGlnValGlglyleuPro 288
```

Alignment Scores:	
Pred. No.:	4,89e-30
Score:	563.50
Percent Similarity:	61.85%
Best Local Similarity:	44.18%
Query Match:	38.65%
DB:	9
	Gaps:
	4

```

Db      230  GlycylproleValcysLeuValGlyGlnSerTrpLeuGlnAlaGlyValIleSerTrp 249
Oy      661  GCGGAGGCGTGTGCCACACCCGACCGGCGCTGGCATTCACACCGCTGTACCTACTGTTG 720
Db      250  GlyGlnGlycysAlaArgGlnAsnArgProGlyValTyrIleArgValThrAlaHisHis 269
Oy      721  GACTGGATCCACCACTATGTCTCCCAA 747
Db      270  AsnTrpIleHisArgIleIleIleProLys 278

RESULT 4
US-10-028-072-222
/ Sequence 222, Application US/10028072
/ Publication No. US20030004311A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Beresini, Maureen
/ APPLICANT: DeForge, Laura
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Filvaroif, Ellen
/ APPLICANT: Gao, Wei-Qiang
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Sherwood, Steven
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Collin K
/ APPLICANT: Wood, William
/ APPLICANT: Zhang
/ TITLE OF INVENTION:
/ FILE REFERENCE:
/ CURRENT APPLICATION NUMBER: US/10/028,072
/ CURRENT FILING DATE: 2001-12-19
/ PRIOR APPLICATION NUMBER: 60/049911
/ PRIOR FILING DATE: 1997-06-18
/ PRIOR APPLICATION NUMBER: 60/056974
/ PRIOR FILING DATE: 1997-08-26
/ PRIOR APPLICATION NUMBER: 60/059113
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059115
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059117
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059122
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059184
/ PRIOR FILING DATE: 1997-09-17
/ PRIOR APPLICATION NUMBER: 60/059263
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059352
/ PRIOR FILING DATE: 1997-09-19
/ PRIOR APPLICATION NUMBER: 60/059588
/ PRIOR FILING DATE: 1997-09-19
/ PRIOR APPLICATION NUMBER: 60/059836
/ PRIOR FILING DATE: 1997-09-24
/ PRIOR APPLICATION NUMBER: 60/062250
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/062285
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/062287
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/062814
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/062816
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063045
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063082
/ PRIOR FILING DATE: 1997-10-31
/

```

PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063561
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063704
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063735
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072320
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818

PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086430
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 19/98-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090538
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

Alignment Scores:
Pred. No.: 4.89e-30
Score: 563.50
Percent Similarity: 61.85%
Best Local Similarity: 44.18%

Length: 290
Matches: 110
Conservative: 44
Mismatches: 86

Query Match: 38.65% Indels: 9
 DB: 9 Gaps: 4
 US-09-598-982-20 (1-771) x US-10-028-072-222 (1-290)

```

QY 16 AGAATGTCGGGGGTGACGAGGCCCCAGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGA 75
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 34 ArgmetValGlyGlnAspThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 53
QY 76 GTCCAGGCGCCCTACTGATGACATTCTGCGGGGGCTCCCTATCCACCCCAAGTGGG 135
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 54 ArgasnGly-----SerHisPheCysGlyGlySerLeuIleAlaGlnGlnTrpVal 70
QY 136 CTGACCGCGCCGCGGTGCTGGGAGCCGCGGCAAGATCTGACCCCTCCAGGGGCA 195
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 71 LeuThrAlaAlaHisCysPhe---ArgAsnThrSerGlnThrSerLeuTrgInValLeu 89
QY 196 CTGCGGGAGCACACCTCTACTAC-----CAGGACCACTGCTGGCGGTCAAGAG 246
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 90 LeuGlyAlaArgGlnLeuValGlnProGlyProHisAlaMetGlyAlaArgValArgGln 109
QY 247 ATCATGTCGACCCCAAGTGTCTACACCGCCCAAGTGGAGCGGACATGCCCTGCTGGAG 306
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 110 ValGluSerAsnProLeuTrgInGlnGlnThrAlaSerSerAlaAspValAlaLeuValGln 129
QY 307 CTGAGAGGACCGGTGAGAGGTCTCAAGCCAGTCCACAGCGTCAACCTGCGCCCTGCTCA 366
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 130 LeuGlnAlaProValProPheThrAsnTrgIleLeuProValCysLeuProAsProSer 149
QY 367 GAGACCTTCCCGCGGGGATGCGCGGTGCTGCTGCTGCTGCGGGCGAGTGGACAATGAT 426
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 150 ValIlePheGlnThrGlyMetAsnCysTrpValThrGlyTrpGlySerProSerGlnGln 169
QY 427 GAGCGCCTCCCAACCCCATTTCTCTGAGAGCAGGTGAAGTCCCATATGCAAAACCC 486
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 170 AspLeuLeuProGlnProArgIleLeuGlnLysLeuAlaValProIleIleAspThrPro 189
QY 487 ATTGTGACGCAAAATACCACTTGGCGGCTACACGCGAGAGACGAGTCCGATCGCGCT 546
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 190 LysCysAsnLeuLeuTrgIleSerLysAspThrGlnPheGlnTrgInProLysThrIleLys 209
QY 547 GACGACATGCTGTGTGCCGG-----AACCCCGAGGAGGACTCATGCCAGGCGCATCC 600
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 210 AsnAspMetLeuCysAlaGlnPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 229
QY 601 GGAGGGCCCCCTGTGTGCAAGTGAATGCGACTGCTCCAGCGCGGCTGTGCTGCTG 660
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 230 GlyGlyProLeuValCysLeuValGlyGlnSerTrpLeuGlnAlaGlyAlaIleSerTrp 249
QY 661 GGCGAGGGCTGTGCCAGCCCAACGCGCTGACATCCCGGTGACTACTACTG 720
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 250 GlyGlnGlyCysAlaArgGlnAsnArgProGlyValTrgIleArgValThrAlaHisHis 269
QY 721 GACTGGATCCACCACTATGTCCCAAA 747
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 270 AsnTrpIleHisArgIleLeuProLys 278
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

LENGTH: 290
 TYPE: PRF
 ORGANISM: Homo sapiens
 US-10-040-655-7

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
4.89e-30	563.50	61.85%	44.18%	38.65%	290	110	44	86	9	4

US-09-598-982-20 (1-771) x US-10-040-655-7 (1-290)

```

QY 16 AGAATGTCGGGGGTGACGAGGCCCCAGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGA 75
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 34 ArgmetValGlyGlnAspThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 53
QY 76 GTCCAGGCGCCCTACTGATGACATTCTGCGGGGGCTCCCTATCCACCCCAAGTGGG 135
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 54 ArgasnGly-----SerHisPheCysGlyGlySerLeuIleAlaGlnGlnTrpVal 70
QY 136 CTGACCGCGCCGCGGTGCTGGGAGCCGCGGCAAGATCTGACCCCTCCAGGGGCA 195
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 71 LeuThrAlaAlaHisCysPhe---ArgAsnThrSerGlnThrSerLeuTrgInValLeu 89
QY 196 CTGCGGGAGCACACCTCTACTAC-----CAGGACCACTGCTGGCGGTCAAGAG 246
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 90 LeuGlyAlaArgGlnLeuValGlnProGlyProHisAlaMetGlyAlaArgValArgGln 109
QY 247 ATCATGTCGACCCCAAGTGTCTACACCGCCCAAGTGGAGCGGACATGCCCTGCTGGAG 306
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 110 ValGluSerAsnProLeuTrgInGlnGlnThrAlaSerSerAlaAspValAlaLeuValGln 129
QY 307 CTGAGAGGACCGGTGAGAGGTCTCAAGCCAGTCCACAGCGTCAACCTGCGCCCTGCTCA 366
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 130 LeuGlnAlaProValProPheThrAsnTrgIleLeuProValCysLeuProAsProSer 149
QY 367 GAGACCTTCCCGCGGGGATGCGCGGTGCTGCTGCTGCTGCGGGCGAGTGGACAATGAT 426
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 150 ValIlePheGlnThrGlyMetAsnCysTrpValThrGlyTrpGlySerProSerGlnGln 169
QY 427 GAGCGCCTCCCAACCCCATTTCTCTGAGAGCAGGTGAAGTCCCATATGCAAAACCC 486
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 170 AspLeuLeuProGlnProArgIleLeuGlnLysLeuAlaValProIleIleAspThrPro 189
QY 487 ATTGTGACGCAAAATACCACTTGGCGGCTACACGCGAGAGACGAGTCCGATCGCGCT 546
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 190 LysCysAsnLeuLeuTrgIleSerLysAspThrGlnPheGlnTrgInProLysThrIleLys 209
QY 547 GACGACATGCTGTGTGCCGG-----AACCCCGAGGAGGACTCATGCCAGGCGCATCC 600
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 210 AsnAspMetLeuCysAlaGlnPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 229
QY 601 GGAGGGCCCCCTGTGTGCAAGTGAATGCGACTGCTCCAGCGCGGCTGTGCTGCTG 660
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 230 GlyGlyProLeuValCysLeuValGlyGlnSerTrpLeuGlnAlaGlyAlaIleSerTrp 249
QY 661 GGCGAGGGCTGTGCCAGCCCAACGCGCTGACATCCCGGTGACTACTACTG 720
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 250 GlyGlnGlyCysAlaArgGlnAsnArgProGlyValTrgIleArgValThrAlaHisHis 269
QY 721 GACTGGATCCACCACTATGTCCCAAA 747
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 270 AsnTrpIleHisArgIleLeuProLys 278
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 6
 US-10-041-006A-9
 Sequence 9, Application US/10041006A
 Patent No. US20020168754A1
 GENERAL INFORMATION:
 APPLICANT: Andrade-Gordon, Patricia

```

; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; TITLE OF INVENTION: Protease T
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/10/041,006A
; CURRENT FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion Protein
; OTHER INFORMATION: of Protease T in a zymogen activation construct
US-10-041-006A-9

Alignment Scores:
Pred. No.: 1,9e-29 Length: 315
Score: 554.50 Matches: 109
Percent Similarity: 60.96% Conservative: 44
Best Local Similarity: 43.43% Mismatches: 89
Query Match: 38.03% Indels: 9
DB: Gaps: 4

US-09-598-982-20 (1-771) x US-10-041-006A-9 (1-315)
OY 10 GAGAAAGAGTGTGGGGGTGAGAGGCCCCAGAGACAGTGGCCCTGAGGTGAGC 69
   ::::::::::::::::::::
DB 49 AspaSprlylIleValIglYlTyralaleuGluglulTrpProtrpIlnValSer 68
OY 70 CTGAGAGTCCAGGCGGCATCTGATGATGACCTTGGGGGGGCGCTTATCCACCCCGAG 129
   ::::::::::::::::::::
DB 69 IleglnArgasnGly-----SerHisPheCysGlyGlySerleuIleaglulgn 85
OY 130 TGGGTGCTGACCGCGCGCGCTGCTGGGAGCGAGCTCAAGATCTGGCGCCCTCAGG 189
   |||||
DB 86 TrpValIleuThrAlaAlaHisCysPhe---ArgAsnThrSerGlnThrSerleuTyrgln 104
OY 190 GTGCACTGCGGGAGACAGCACTTACTAC-----CAGACCAGCTGCTGCCGGTC 240
   |||
DB 105 ValIleuLeuGlyAlaArgIleuValIleuProgluProHisAlaMetTyralaArgVal 124
OY 241 AGCAGATCATGTCACCCAGTTCATACCGCCAGATGGAGCGGACATCGCCCTG 300
   ::::::::::::::::::::
DB 125 ArgGlnValIleuSerAsnProleuTyrglnIleuThrAlaSerSerAlaAspValAlaIleu 144
OY 301 CTGAGAGTCCAGGAGCGCGGTGAAGGTCTCCAGACGATCCACGATCCACCTGCCCT 360
   ::::::::::::::::::::
DB 145 ValGlnleuGlnAlaProValProPheThrAsnTyrlleuProValLysIleuProasp 164
OY 361 GCCTAGAGACCTTCCCCCGGGGATGCCGTGCTGGGTACTGCTGGGGCGCATGTGAC 420
   |||
DB 165 ProSerValIlePhehegluThrIleuMetasnCysTrpValThrGlyTrpIleuSerProser 184
OY 421 AATGTAGACCGCTCCACCGCATTTCTCTGAAGCAGAGTGAAGTCCCAATATGAA 480
   ::::::::::::::::::::
DB 185 GluGlnAspIleuLeuProgluProArgIleuGlnIleuValAlaValProIleuAsp 204
OY 481 AACCAATTTGTGACGAAATATACCACTTGGCGCTACACGAGACAGCATCCGCATC 540
   |||
DB 205 ThrProLysCysAsnIleuLeuTyrlSerLysAspThrIleuPheGlyTyrlGlnProLysThr 224
OY 541 GTCCGTGAGACATCTGTGTGCCGG-----AACCCCGAGGAGACTCATGCCAGGC 594
   ::::::::::::::::::::
DB 225 IleLysAsnAspMetIleuLysIleuGlyPhegluGlnIleuLysAspIleuLysGly 244
OY 595 GACCTCGGAGGCGCGCTGCTGTCGAAGGATGGACACTGGCTCAGCGCGCGCTGGTC 654
   |||
DB 245 AspSerIleuGlyIleuLeuValLysIleuValIleuGlnSerTrpIleuGlnIleuValIle 264
OY 655 AGCTGGGCGAGGCTGTGCCAGCCCAACCGCGCTGGCATCTACACCGTGCACCTAC 714
```

```

DB 265 SerTrpGlyGluGlyCysAlaArgIleuAsnArgProgluValTyrlleuArgValThrAla 284
OY 715 TACTTGAGTGCATCCACCATATGTCCCAAA 747
   ::::::::::::::::::::
DB 285 HlshAsnTrpIleHisArgIleuProLys 295

RESULT 7
US-10-040-655-9
; Sequence 9, Application US/10040655
; Patent No. US20020146805A1
; GENERAL INFORMATION:
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; TITLE OF INVENTION: Protease T
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/10/040,655
; CURRENT FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion Protein
; OTHER INFORMATION: of Protease T in a zymogen activation construct
US-10-040-655-9

Alignment Scores:
Pred. No.: 1,9e-29 Length: 315
Score: 554.50 Matches: 109
Percent Similarity: 60.96% Conservative: 44
Best Local Similarity: 43.43% Mismatches: 89
Query Match: 38.03% Indels: 9
DB: Gaps: 4

US-09-598-982-20 (1-771) x US-10-040-655-9 (1-315)
OY 10 GAGAAAGAGTGTGGGGGTGAGAGGCCCCAGAGACAGTGGCCCTGAGGTGAGC 69
   ::::::::::::::::::::
DB 49 AspaSprlylIleValIglYlTyralaleuGluglulTrpProtrpIlnValSer 68
OY 70 CTGAGAGTCCAGGCGGCATCTGATGATGACCTTGGGGGGGCGCTTATCCACCCCGAG 129
   ::::::::::::::::::::
DB 69 IleglnArgasnGly-----SerHisPheCysGlyGlySerleuIleaglulgn 85
OY 130 TGGGTGCTGACCGCGCGCGCTGCTGGGAGCGAGCTCAAGATCTGGCGCCCTCAGG 189
   |||||
DB 86 TrpValIleuThrAlaAlaHisCysPhe---ArgAsnThrSerGlnThrSerleuTyrgln 104
OY 190 GTGCACTGCGGGAGACAGCACTTACTAC-----CAGACCAGCTGCTGCCGGTC 240
   |||
DB 105 ValIleuLeuGlyAlaArgIleuValIleuProgluProHisAlaMetTyralaArgVal 124
OY 241 AGCAGATCATGTCACCCAGTTCATACCGCCAGATGGAGCGGACATCGCCCTG 300
   ::::::::::::::::::::
DB 125 ArgGlnValIleuSerAsnProleuTyrglnIleuThrAlaSerSerAlaAspValAlaIleu 144
OY 301 CTGAGAGTCCAGGAGCGCGGTGAAGGTCTCCAGACGATCCACGATCCACCTGCCCT 360
   ::::::::::::::::::::
DB 145 ValGlnleuGlnAlaProValProPheThrAsnTyrlleuProValLysIleuProasp 164
OY 361 GCCTAGAGACCTTCCCCCGGGGATGCCGTGCTGGGTACTGCTGGGGCGCATGTGAC 420
   |||
DB 165 ProSerValIlePhehegluThrIleuMetasnCysTrpValThrGlyTrpIleuSerProser 184
OY 421 AATGTAGACCGCTCCACCGCATTTCTCTGAAGCAGAGTGAAGTCCCAATATGAA 480
   ::::::::::::::::::::
DB 185 GluGlnAspIleuLeuProgluProArgIleuGlnIleuValAlaValProIleuAsp 204
```

```
QY 481 AACCAATTGTGACGCCAAATACCACTTGGCGCCTACACGGAGACGACGTCGCCATC 540
    |||::: |||
Db 205 ThrProLysCysAsnLeuLeuTyrSerLysAspThrGlnPheGlyTyrGlnProLysThr 224
QY 541 GTCGGGAGACATGTCGTGTGCCGGG-----AACACCCGGAGGACATGACCCAGGGC 594
    :::::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 225 lIeLysAsnAspMetLeuLysAlaGlyPheGlnGlyLysLysAspAlaGlyLysGly 244
QY 595 GACTCCGGAGGGCCCCGTGTGTGCAAGTGGAATGGCAGCTGGCTGCAGCGGGCGTGC 654
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 245 AspSerGlyGlyProLeuValAlcysLeuValGlyGlnSerTrpLeuGlnAlaGlyVal 264
QY 655 ACCTGGGGGAGGGCTGTGTCGCCAGCCCAACCGGCTGCATGCATACCCGTCACCTAC 714
    |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 265 SerTrpGlyGlyGlyCysAlaArgGlnAsnArgProGlyValTyrTrlLeaArgValThrAla 284
QY 715 TACTTGAGTGCATCCACCATATGTCGCCAA 747
    :::::|||||:::|||||
Db 285 HlshIsAsnTrpIleHlshArgIleIleProLys 295

RESULT 8
US-10-041-400A-7
: Sequence 7, Application US/10041400A
: Patent No. US20020110895A1
: GENERAL INFORMATION:
: APPLICANT: Darlow, Andrew
: APPLICANT: Andrade-Gordon, Patricia
: APPLICANT: Qi, Jensen
: TITLE OF INVENTION: DNA Encoding the Human Serine
: TITLE OF INVENTION: Protease EOS
: FILE REFERENCE: ORT-1031
: CURRENT APPLICATION NUMBER: US/10/041,400A
: CURRENT FILING DATE: 2002-01-08
: PRIOR APPLICATION NUMBER: US/09/387,375
: PRIOR FILING DATE: 1999-08-31
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 7
: LENGTH: 284
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-041-400A-7

Alignment Scores:
Pred. No.: 2,866-28      Length: 284
Score: 536.50           Matches: 112
Percent Similarity: 58.54%      Conservative: 32
Best Local Similarity: 45.53%      Mismatches: 93
Query Match: 36.80%           Indels: 9
DB: 12                  Gaps: 4

US-09-598-982-20 (1-771) x US-10-041-400A-7 (1-284)
QY 7 CTCGAGAAAGAATGTCGCGGGGTGAGAGGCCCCAGAGGACGAGTGGCCAGAGTG 66
    :::::|||||:::|||||:::|||||
Db 33 MetSerSerArgIleValAlcGlyAlaArgAspGlyValGlnTrpProThrAla 52
QY 67 ACCCTGAGAGTCCACGGCCCATACTGATGCATCTTGGGGGGGCTCCCTCATCCAGCC 126
    |||:::|||||:::|||||:::|||||
Db 53 SerIleGlnHisProGly-----AlaHisValCysGlyGlySerLeuIleAlaPro 69
QY 127 CAGTGGGTGCTGACCGCGCGGCTGTCGGAGCGGAGCTGACAGGATGCTGGCGGCGCTC 186
    |||:::|||||:::|||||
Db 70 GlnTrpValLeuThrAlaAlaHisCysPhe---ProArgAlaGlnAlaLeuProAlaGlnTyr 88
QY 187 AGGGTCAACTGTCGGGAGAGGACACTC-----TACTACAGGAGCAGGACGTCGTCGG 237
    |||:::|||||:::|||||
Db 89 ArgValAlaArgLeuGlyAlaLeuAlaArgLeuGlySerThrSerProArgThrLeuSerValPro 108
QY 238 GTCAAGAGAGATGATGTCGACCCACAGTTCTACACGCGCCCAAGATGGAGCGAGCATGCGC 297
    |||:::|||||:::|||||
Db 109 ValArgAlaValLeuLeuLeuProAspTyrSerGlnAspGlyAlaAlaArgLysAspLeuAla 128
```

```
QY 298 CTCGTGAGACCTGAGAGAGCGGCTGAAGTCTCCAGACGATCCACAGGTCACCTGCC 357
    |||:::|||||:::|||||:::|||||
Db 129 LeuLeuGlnLeuAlaArgProValProLeuSerAlaArgValGlnProValAlcLysLeuPro 148
QY 358 CCTGCTCAGAAACCTTCCCGCGGAGATGCCGTGGTGTACTGTGCTGCGGGCCATGTCG 417
    |||:::|||||:::|||||:::|||||
Db 149 ValProGlyAlaArgProProProGlyThrProCysArgValThrGlyTrpIleSerLeu 168
QY 418 GACATATGATGAGCGGCTCCACCGGCATTTGCTCGAAGACAGCAAGGATGAGGCTGAATG 477
    |||:::|||||:::|||||:::|||||
Db 169 ArgProGlyValProLeuProGlnTrpArgProLeuGlnGlyValAlaArgValProLeuLeu 188
QY 478 GAACAACCAATTTGTGACGCCAAATACCACTTGGCGCCTACACGGAGAGACGACGTCGC 537
    :::::|||||:::|||||:::|||||
Db 189 AspSerArgTrhCysAspGlyLeuTyrHisValGlyAlaAspValProGlnAlaGlyArg 208
QY 538 ATGTCGCTGAGACACATGCTGTGTGCCGGG-----AACACCCGGAGGAGACTCATGCGCAG 591
    |||:::|||||:::|||||
Db 209 lIeValLeuProGlySerLeuLysAlaGlyTyrProGlnGlyHisLysAspAlaCysGln 228
QY 592 GCGGACTCCGGAGGGCCCTGTGTGTCAGAGTGGAATGGCAGCCTGGCTGCAGCGGGCGCTG 651
    |||:::|||||:::|||||:::|||||
Db 229 GlYAspSerGlyLysProLeuThrCysLeuGlnSerGlySerTrpValLeuValGlyVal 248
QY 652 CTCAGCTGGGCGGAGGCTGTGTCGCCAGCCCAACCGGCTGCATGCATACCCGTCATACC 711
    |||:::|||||:::|||||:::|||||
Db 249 ValSerTrpGlyLysGlyCysAlaLeuProAsnArgProGlyValTyrTrhSerValAla 268
QY 712 TACTTACTTGACTGCATC 729
    |||
Db 269 ThrTyrSerProThrIle 274

RESULT 9
US-10-041-264A-7
: Sequence 7, Application US/10041264A
: Patent No. US20020142446A1
: GENERAL INFORMATION:
: APPLICANT: Darlow, Andrew
: APPLICANT: Andrade-Gordon, Patricia
: APPLICANT: Qi, Jensen
: TITLE OF INVENTION: DNA Encoding the Human Serine
: TITLE OF INVENTION: Protease EOS
: FILE REFERENCE: ORT-1031
: CURRENT APPLICATION NUMBER: US/10/041,264A
: CURRENT FILING DATE: 2002-01-08
: PRIOR APPLICATION NUMBER: US/09/387,375
: PRIOR FILING DATE: 1999-08-31
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 7
: LENGTH: 284
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-041-264A-7

Alignment Scores:
Pred. No.: 2,866-28      Length: 284
Score: 536.50           Matches: 112
Percent Similarity: 58.54%      Conservative: 32
Best Local Similarity: 45.53%      Mismatches: 93
Query Match: 36.80%           Indels: 9
DB: 12                  Gaps: 4

US-09-598-982-20 (1-771) x US-10-041-264A-7 (1-284)
QY 7 CTCGAGAAAGAATGTCGCGGGGTGAGAGGCCCCAGAGGACGAGTGGCCAGAGTG 66
    :::::|||||:::|||||:::|||||
Db 33 MetSerSerArgIleValAlcGlyAlaArgAspGlyValGlnTrpProThrAla 52
QY 67 ACCCTGAGAGTCCACGGCCCATACTGATGCATCTTGGGGGGGCTCCCTCATCCAGCC 126
    |||:::|||||:::|||||:::|||||
Db 53 SerIleGlnHisProGly-----AlaHisValCysGlyGlySerLeuIleAlaPro 69
```

```
OY 127 CAGTGGTCTGACCGCGCGGCTGCTGGAGCGAGCATCAAGATCGCCGCCCTC 186
Db 70 GINTRPVallleuThralalaHisCysphe---ProArgArgAlaLeuProAlaGluTr 88
OY 187 AGGTGCAACTCGCGGAGCAGCAGCTC-----TACTACAGGAGCAGCTGCGCCG 237
Db 89 ArgValArgLeuGlyAlaLeuArgLeuGlySerThrSerProArgThrLeuSerValPro 108
OY 238 GTACACAGGATATCGTGCACCCAGATTCTACACCGCCCAATGGAGCGGCATCGCC 297
Db 109 ValArgValLeuLeuProProAspPylSerGluAspGlyAlaArgGlyAspLeuAla 128
OY 288 CTGCTGAGCTGAGAGAGCGCGGTGAGCTCTCACACCACTCCACAGCTCACCTGCCC 357
Db 129 LeuLeuGlnLeuArgArgProValProLeuSerAlaArgValGlnProValCysLeuPro 148
OY 338 CCTGCTCAGAGACCTTCCCCCGGGGATGCCGTGCTGCTGCTGCTGCTGCTGCTG 417
Db 149 ValProGlyAlaArgProProProGlyThrProCysArgValThrGlyTrpGlySerLeu 168
OY 418 GACATGATGAGCGCGCTCCACCGCATTTCTCTGAGAGCAGTGAAGTCCCATTAATG 477
Db 169 ArgProGlyValProLeuProGluTrpArgProLeuGlnGlyValArgValProLeuLeu 188
OY 478 GAAACACACATTTGTAGCCAAATACACCTTGGCGCCTACACGGAGACGAGCTCCGC 537
Db 189 AspSerArgThrCysAspPylLeuTrpHisValGlyAlaAspValProGlnAlaGluArg 208
OY 538 ATGCTCCGTGACGACATGCTGTGTGCGCGG-----AACACCCGGAGGACTCATGCCAG 591
Db 209 IleValLeuProGlySerLeuCysAlaGlyTyrProGlnGlnHisLysAspAlaCysGln 228
OY 592 GCGGACTCCGGAGGCGCCCTGCTGTCGAAGTGAATGGACCTGGCTGAGCGGCGCTG 651
Db 228 GlyAspSerGlyGlyProLeuThrCysLeuGlnSerGlySerThrValLeuValGlyVal 248
OY 652 GTCACTGGGGGAGGGCTGTGCTCCAGCCCAACGCGCTGCATGTACACCGCTGCACC 711
Db 249 ValSerTrpGlyLysGlyCysAlaLeuProAsnArgProGlyValTyrThrSerValAla 268
OY 712 TACTACTTGACTGGATC 729
Db 269 ThrTyrSerProTrpIle 274

RESULT 10
US-10-042-091A-7
; Sequence 7, Application US/10042091A
; Patent No. US20020142447A1
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Qi, Jensen
; TITLE OF INVENTION: DNA Encoding the Human Serine
; FILE REFERENCE: ORT-1031
; CURRENT APPLICATION NUMBER: US/10/042,091A
; CURRENT FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US/09/387,375
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-091A-7

Alignment Scores:
Pred. No.: 2,86e-28 Length: 284
Score: 536.50 Matches: 112
Percent Similarity: 58.54% Conservative: 32
Best Local Similarity: 45.53% Mismatches: 93
Query Match: 36.80% Indels: 9
```

```
DB: 12 Gaps: 4
US-09-598-982-20 (1-771) x US-10-042-091A-7 (1-284)
OY 7 CTCGAAAAGACATCTGTGGGGGTACAGAGCGCCCGACAGCACTGGCCCTGCAGCTG 66
Db 33 MetSerSerArgIleValGlyArgAspGlyArgAspGlyGluTrpProTrpAla 52
OY 67 AGCGTACAGCTCCACCGCCATACGTGACATCTTCCTGGGGGCTCCCTCATCCACCC 126
Db 53 SerIleGlnHisProGly-----AlaHisValCysGlyGlySerLeuLeuAlaPro 69
OY 127 CAGTGGTCTGACCGCGCGGCTGCTGGAGCGAGCATCAAGATCGCCGCCCTC 186
Db 70 GINTRPVallleuThralalaHisCysphe---ProArgArgAlaLeuProAlaGluTr 88
OY 187 AGGTGCAACTCGCGGAGCAGCAGCTC-----TACTACAGGAGCAGCTGCGCCG 237
Db 89 ArgValArgLeuGlyAlaLeuArgLeuGlySerThrSerProArgThrLeuSerValPro 108
OY 238 GTACACAGGATATCGTGCACCCAGATTCTACACCGCCCAATGGAGCGGCATCGCC 297
Db 109 ValArgValLeuLeuProProAspPylSerGluAspGlyAlaArgGlyAspLeuAla 128
OY 288 CTGCTGAGCTGAGAGAGCGCGGTGAGCTCTCACACCACTCCACAGCTCACCTGCCC 357
Db 129 LeuLeuGlnLeuArgArgProValProLeuSerAlaArgValGlnProValCysLeuPro 148
OY 338 CCTGCTCAGAGACCTTCCCCCGGGGATGCCGTGCTGCTGCTGCTGCTGCTGCTG 417
Db 149 ValProGlyAlaArgProProProGlyThrProCysArgValThrGlyTyrGlySerLeu 168
OY 418 GACATGATGAGCGCGCTCCACCGCATTTCTCTGAGAGCAGTGAAGTCCCATTAATG 477
Db 169 ArgProGlyValProLeuProGluTrpArgProLeuGlnGlyValArgValProLeuLeu 188
OY 478 GAAACACACATTTGTAGCCAAATACACCTTGGCGCCTACACGGAGACGAGCTCCGC 537
Db 189 AspSerArgThrCysAspPylLeuTrpHisValGlyAlaAspValProGlnAlaGluArg 208
OY 538 ATGCTCCGTGACGACATGCTGTGTGCGCGG-----AACACCCGGAGGACTCATGCCAG 591
Db 209 IleValLeuProGlySerLeuCysAlaGlyTyrProGlnGlnHisLysAspAlaCysGln 228
OY 592 GCGGACTCCGGAGGCGCCCTGCTGTCGAAGTGAATGGACCTGGCTGAGCGGCGCTG 651
Db 228 GlyAspSerGlyGlyProLeuThrCysLeuGlnSerGlySerThrValLeuValGlyVal 248
OY 652 GTCACTGGGGGAGGGCTGTGCTCCAGCCCAACGCGCTGCATGTACACCGCTGCACC 711
Db 249 ValSerTrpGlyLysGlyCysAlaLeuProAsnArgProGlyValTyrThrSerValAla 268
OY 712 TACTACTTGACTGGATC 729
Db 269 ThrTyrSerProTrpIle 274

RESULT 11
US-10-041-400A-9
; Sequence 9, Application US/10041400A
; Patent No. US20020110895A1
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Qi, Jensen
; TITLE OF INVENTION: DNA Encoding the Human Serine
; FILE REFERENCE: ORT-1031
; CURRENT APPLICATION NUMBER: US/10/041,400A
; CURRENT FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US/09/387,375
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
```

```

: SEQ ID NO 9
:
: LENGTH: 316
:
: TYPE: prt
:
: ORGANISM: Artificial Sequence
:
: FEATURE:
:
: OTHER INFORMATION: Description of Artificial Sequence: Amino acid
:
: OTHER INFORMATION: sequence of Eos zymogen fusion gene
:
: OS-10-041-400A-9

```

Alignment Scores:	
Pred. No.:	1,11e-27
Score:	52.50
Percent Similarity:	57.96%
Best Local Similarity:	45.31%
Query Match:	36.18%
DB:	12
Length:	316
Matches:	111
Conservative:	31
Mismatches:	94
Indels:	3
Gaps:	4

US-09-598-982-20 (1-771) x US-10-041-400A-9 (1-316)

[illegible]

```

1      : Sequence 9, Application US/10041264A
2      : Patent No. US20020142446a1
3      : GENERAL INFORMATION:
4      : APPLICANT: Darrow, Andrew
5      : APPLICANT: Andrade-cordon, Patricia
6      : APPLICANT: Qi, Jensen
7      : TITLE OF INVENTION: DNA Encoding the Human Serine
8      : TITLE OF INVENTION: Protease EOS
9      : FILE REFERENCE: CRT-1031
10     : CURRENT APPLICATION NUMBER: US/10/041,264A
11     : CURRENT FILING DATE: 2002-01-08
12     : PRIOR APPLICATION NUMBER: US/09/387,375
13     : PRIOR FILING DATE: 1999-08-31
14     : NUMBER OF SEQ ID NOS: 9
15     : SOFTWARE: PatentIn Ver. 2.0
16     : SEQ ID NO 9
17     :
18     : LENGTH: 316
19     :
20     : TYPE: PRT
21     : ORGANISM: Artificial Sequence
22     : FEATURE:
23     : OTHER INFORMATION: Description of Amino acid
24     : OTHER INFORMATION: sequence of EOS zymogen fusion gene
25     : US-10-041-264A-9

```

Alignment Scores:	
Pred. No.:	1 11e-27
Score:	527.50
Percent Similarity:	57.96%
Best Local Similarity:	45.31%
Query Match:	36.18%
DB:	12
	4
	9
	4
	316
	111
	31
	94
	94
	11
	4
	9
	4

US-09-598-982-20 (1-771) X US-10-041-264A-9 (1-316)

QY	10	GAAGAAACAATGCTGGGGGGGTACAGAGGCCCAAGAGCAAGTGGCCCTGGACAGTACG	63
	11	64
Db	49	AspAspLysIleValaIGlyIlytLAlaLeuLuspsglIutPrProTglnAlaIAser	68
QY	70	CTGAGAGTCCACGGCGCCCTACTGGATGCACTTCGTCGGGGGGCTCCGCATCCACCCAG	129
Db	69	85
QY	130	TGGGTGCTGACCGCGCGCGCTGGGTGGAGCGGAGCAAGATTCGGCGCCCTGAG	189
Db	86	TrpValLeuThrAlaAlaHisCysPhe---ProArgAlaAlaLeuProAlaGlytArg	104
QY	190	GTGCAACTGCGGGAGCAACACTC-----TACTACAGAGCAACAGTCTCCGGTC	240
Db	105	244
QY	241	ACGAGATCATGTGTACCCACCAAGTTTCACACCGCCGAGTCGGAGGGAGATCGCCCTG	300
Db	125	144
QY	301	CTGAGACTGAGAGACCGGTGAGAGGTTCACAGCCACGTCACACAGTCCAGCTCCGCCCT	360
Db	145	LeuGlnLeuArgArgProValProLeuSerAlaArgAlaGlnProValCysLeuProVal	164
QY	361	GCCTAGAGACCTTCCCGCGGGAGTCGCGTCTGGGTCACTGGCTGGGGCGATGTGGAC	420
Db	165	ProGlyAlaArgProProTglnThrProCysArgValAlaThrGlyTglnSerLeuArg	184
QY	421	AATGATGAGCGCGTCCACCGCATTTCTCTGTAAGCAAGGGAAGGCCCATTAATGGAA	480
Db	185	ProGlyValProLeuProGlnTrpArgProLeuGlnGlyAlaArgValProLeuLeuASP	204
QY	481	AACCAATTGTGACAGCAAAATACCACTTGGCGCTTACAGGGAGAGCAAGTCCGATC	540
Db	205	SerArgThrCysAspGlyLeuThrIstAlaAlaLysPvalProGlnAlaGlnArgIle	224
QY	541	GTCCTGACGACATCTCTGTCCGGG-----AACACCGGAGGAGCTATGCCAGGCC	594
Db	225	ValLeuProGlySerLeuCysAlaGlyIlytPrProGlnGlnHisIstAspAlaCysGlnGly	244

```
QY 595 GACTCCGAGGAGCCCTGTGTGCAAGGTGAATGGACCTGTGTCAGCGGGCTGTGTC 654
|||||
Db 245 AspserglyglyProleuthrCysleuInserglySertrpValLeuValGlyValVal 264
QY 655 AGCTGGGCGAGGGCTGTGCCACGCCACCGGCTGTGATTCACCCGGTGCACCTAC 714
|||||
Db 265 SertrpGlyglyCysAlaLeuProAsnArgProGlyValTyrThrSerValaThr 284
QY 715 TACTTGAGCTGATC 729
|||
Db 285 TyrSerProtrpIle 289

RESULT 13
US-10-042-091A-9
; Sequence 9, Application US/10042091A
; Patent No. US20020142447A1
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Qi, Jensen
; TITLE OF INVENTION: DNA Encoding the Human Serine
; TITLE OF INVENTION: Protease EOS
; FILE REFERENCE: ORT-1031
; CURRENT APPLICATION NUMBER: US/10/042,091A
; CURRENT FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US/09/387,375
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid
; OTHER INFORMATION: Sequence of EOS zymogen fusion gene
US-10-042-091A-9

Alignment Scores:
Pred. No.: 1,11e-27 Length: 316
Score: 527.50 Matches: 111
Percent Similarity: 57.96% Conservative: 31
Best Local Similarity: 45.31% Mismatches: 94
Query Match: 36.18% Indels: 9
Gaps: 4

US-09-598-982-20 (1-771) x US-10-042-091A-9 (1-316)
QY 10 GAGAAAGAAATCGTCGGGGGTGTCAGAGAGCCGCCAGAGACAGTGGCCCTGCAGCTGAC 69
::: |||||
Db 49 AspAspIysIleValGlyGlyTyrAlaLeuGluAspGlyIuTrpProTrpLeuAlaSer 68
QY 70 CTGAGAGTCCACGGCCCACTACTGATGCATCTTGCGGGGCTCCCTATCCACCCCAAG 129
::: |||
Db 69 IlegInHsPrgGly-----AlaHsValCysGlySerLeuIleAlaProGln 85
QY 130 TGGGTCGTGACCGGCGGCGGTGCGTGGAGCCGACGTCAAGATCTGCGCCCTCAGG 189
|||||
Db 86 TrpValLeuThrAlaAlaHsCysPhe---ProArgAlaAlaLeuProAlaGluTyrArg 104
QY 190 GTGCAGACTGCGGAGCAGACACTC-----TACTACAGAGACAGACTCTGCGGGTGC 240
|||||
Db 105 ValArgLeuGlyAlaLeuArgLeuGlySerThrSerProArgThrLeuSerValProVal 124
QY 241 AGCAGAGTCACTGTCACCCACAGTTTACACCGCCGACATCGAGCGGACATCGCCCTG 300
|||||
Db 125 ArgArgValLeuLeuProProAspTyrSerGluAspGlyAlaArgGlyAspLeuAlaLeu 144
QY 301 CTGAGAGTGGAGAGAGCGGTGAGGTCTCCAGCCAGCTCCACAGCGGTACACCTGCCCCCT 360
|||||
Db 145 LeuGlnLeuArgArgProValProLeuSerAlaArgValGlnProValCysLeuProVal 164
```

```
QY 361 GCCTCAGAGACCTTCCCGCGGGAGATGCCGTGTGGTCACTGGCTGGGGCATGTGAC 420
|||||
Db 165 ProGlyAlaArgProProGlyThrProCysArgValThrGlySerLeuArg 184
QY 421 AATGATGAGCCGCTCCACCGCCATTTCTCTGAAGCAGGTGAAAGTCCCATATGAA 480
|||||
Db 185 ProGlyValProLeuProGluTrpArgProLeuGlnGlyValArgValProLeuAsp 204
QY 481 AACCAATTTGTGACGCAAAATACACCTTGCGGCTGACAGCGGAGACAGACGTCGATC 540
::: |||||
Db 205 SerArgThrCysAspGlyLeuThrHisValGlyAlaAspValProGlnAlaGluArgIle 224
QY 541 GTCCGTGACAGACATCTGTGTGCCGG-----AACACCGGAGGAGACTATGCCAGGCG 594
|||
Db 225 ValLeuProGlySerLeuCysAlaGlyTyrProGlnGlyHisAspAlaCysGlnGly 244
QY 595 GACTCCGAGGAGCCCTGTGTGCAAGGTGAATGGACCTGTGTCAGCGGGCTGTGTC 654
|||||
Db 245 AspserglyglyProleuthrCysleuInserglySertrpValLeuValGlyValVal 264
QY 655 AGCTGGGCGAGGGCTGTGCCACGCCACCGGCTGTGATTCACCCGGTGCACCTAC 714
|||||
Db 265 SertrpGlyglyCysAlaLeuProAsnArgProGlyValTyrThrSerValaThr 284
QY 715 TACTTGAGCTGATC 729
|||
Db 285 TyrSerProtrpIle 289

RESULT 14
US-09-948-094-2
; Sequence 2, Application US/09948094
; Patent No. US20020090625A1
; GENERAL INFORMATION:
; APPLICANT: The Brigham and Women's Hospital, Inc.
; APPLICANT: Mok, Samuel
; APPLICANT: Wong, Kwong-kwok
; TITLE OF INVENTION: Methods of Detecting Cancer Based on Prostatin
; FILE REFERENCE: 81994/282423
; CURRENT APPLICATION NUMBER: US/09/948,094
; CURRENT FILING DATE: 2001-09-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-948-094-2

Alignment Scores:
Pred. No.: 1,15e-26 Length: 343
Score: 512.00 Matches: 103
Percent Similarity: 57.03% Conservative: 39
Best Local Similarity: 41.37% Mismatches: 97
Query Match: 35.12% Indels: 10
Gaps: 6

US-09-598-982-20 (1-771) x US-09-948-094-2 (1-343)
QY 10 GAGAAAGAAATCGTCGGGGGTGTCAGAGAGCCGCCAGAGACAGTGGCCCTGCAGCTGAC 69
::: |||||
Db 42 GlnAlaArgIleThrGlyGlySerSerAlaValAlaGlyIuTrpProTrpGlnValSer 61
QY 70 CTGAGAGTCCACGGCCCACTACTGATGCATCTTGCGGGGCTCCCTATCCACCCCAAG 129
::: |||
Db 62 IleThrTyrGluGly-----ValHisValCysGlySerLeuValSerGluGln 78
QY 130 TGGGTCGTGACCGGCGGCGGTGCGTGGAGCCGACGTCAAGATCTGCGCCCTCAGG 189
|||||
Db 79 TrpValLeuSerAlaAlaHsCysPhe---ProSerGlnHisHisLysGlnAlaTyrGly 97
QY 190 GTGCAGACTGCGGAGCAGACACTC-----TACTACAGAGAC---CAGCTGTCTCCGGTGC 240
|||||
Db 98 ValLysLeuGlyAlaHisGlnLeuAspSerTyrSerGluAspAlaLysValSerThrLeu 117
```

```

QY 241 AGCAGATCATGTGTCACCCACAGATTCTTACACCGCCAGATGCGAGCGGACATGCGCCCTG 300
    ||||| ||||| ::: ||||| |||||
Db 118 LysAspIleIleIleProHisProSerTyrIleuGlnIleuGlySerGlnIleValAspIleIleLeu 137
QY 301 CTGAGCTGGAGGAGCGCGTGAAGGTCTCCACGACGTCACGACGATCCACCTGCGCCCT 360
    ||||| ||||| ||||| ||||| |||||
Db 138 LeuGlnIleuSerArgProIleThrPheSerArgTyrIleArgProIleCysLeuProIle 157
QY 361 GCGTCAGACACCTTCCCCCGGGGATGCCGTGCTGACGTGCGTGGGGGCGATGTGAC 420
    ||||| ||||| ||||| ||||| ||||| |||||
Db 158 AlaAsnAlaSerPheProAsnGlyLeuHisCysThrValThrGlyTyrPheIleValAla 177
QY 421 AATGATGAGCCGCTCCACCGCATTTCTCTGAAGCAGTGAAGTCCCATTAATGAA 480
    ||||| ||||| ||||| ||||| ||||| |||||
Db 178 ProSerValSerLeuIleuThrProLysProLeuGlnIleuGlnIleValProLeuIleSer 197
QY 481 AACCAATTTGTGACGCAAAATACACACTTGGCGCTACACGCGAGACGACGCTCCGCATC 540
    ||||| ||||| ||||| ||||| ||||| |||||
Db 198 ArgGluThrCysAsnGlyLeuTyrAsnIleAspAla---LysProGluIleuProHisPhe 216
QY 541 GTCCCTGAGACATGCTGTGTGTCGCGGAGACACCCG-----AGGACTCATGCCAGGCG 594
    ||||| ||||| ||||| ||||| ||||| |||||
Db 217 ValGlnIleuAspMetValCysAlaGlyTyrValIleuGlyGlyLysAspAlaCysGlnGly 236
QY 595 GACTCCGAGGAGCGCCCTGTGTGCAAGGTGAATGGCAGCTGGCTGACGCGGCGGTGTC 654
    ||||| ||||| ||||| ||||| ||||| |||||
Db 237 AspSerGlyGlyProLeuSerCysProValIleuGlyLeuTyrPheIleuThrGlyIleVal 256
QY 655 AGCTGGGCGAGGCGCTGTGCCAGGCCACGCGCTGCATCTACACCGCTGCACCTAC 714
    ||||| ||||| ||||| ||||| ||||| |||||
Db 257 SerTrpGlyAspAlaCysGlyAlaArgAsnArgProGlyValTyrThrIleuAlaSerSer 276
QY 715 TACTTGAGATGGATGCACCACTATGTC 741
    ||||| ||||| ||||| |||||
Db 277 TyrAlaSerTrpIleGlnSerIleVal 285

RESULT 15
US-10-067-761-19
; Sequence 19, Application US/10067761
; Publication No. US2002019701A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: Serine protease Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PTO05P4
; CURRENT APPLICATION NUMBER: US/10/067,761
; CURRENT FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 09/804,156
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/189,025
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-067-761-19

```

Alignment Scores:

Pred. No.:	3,53e-26	Length:	238
Score:	504.50	Matches:	106
Percent Similarity:	59.73%	Conservative:	26
Best Local Similarity:	47.96%	Mismatches:	76
Query Match:	34.60%	Indels:	13
DB:	9	Gaps:	6

US-09-598-982-20 (1-771) x US-10-067-761-19 (1-238)

```

QY 16 AGAATGCTGGGGGTCAGAGGCGCCCGAGAGCAATGCGCCCTGGCAGCTGAGCTGAGA 75
    ||||| ||||| ||||| ||||| ||||| |||||
Db 29 ArgIleValGlyGlyHisAlaIleProValGlyAlaIleTrpProTrpGlnAlaSerLeuArg 48

```

```

QY 76 GTCCAGCGCCCATTAAGTGCATTTCTGCGGGGCTCCCTCATCCACCCAGATGGGTG 135
    ::: ||||| ||||| ||||| ||||| ||||| |||||
Db 49 LeuArgArg-----ValHisValCysGlyGlySerLeuIleuSerProIleTrpVal 65
QY 136 CTGACCGCGCGCGGTGCGTGGGACCGGAGCTCAAGATGTGCGCGCTCAGGTTGCA 195
    ||||| ||||| ||||| ||||| ||||| |||||
Db 66 LeuThrAlaIleHisCysPheSerGlySerLeuAsnSer---SerAspTrpGlnValHis 84
QY 196 CTGCGGAGACACACCTGTCACTACAGAGACAGCTGCTCCCGGTGACAGATCATCTG 255
    ||||| ||||| ||||| ||||| ||||| |||||
Db 85 LeuGlyGlyLeuGlnIleThrLeuSerProHisPheSerThrValArgIleIleLeu 104
QY 256 CACCCACATTTCTTACACCGCCAGATCGA-----GCGACATGCGCTGCGGAGCTG 309
    ||||| ||||| ||||| ||||| ||||| |||||
Db 105 HisSerSer---ProSerGlyIleProGlyThrSerGlyAspIleAlaIleuValGluLeu 123
QY 310 GAGGAGCGGTGAAGGTCTCCAGCAGCTCCACAGGTACCGTACCGCTCCCGCTCAGAG 369
    ||||| ||||| ||||| ||||| ||||| |||||
Db 124 SerValProValThrLeuSerSerArgIleLeuProValCysLeuProGlyAlaSerAsp 143
QY 370 ACCTTCCCGCGGGGATGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 429
    ||||| ||||| ||||| ||||| ||||| |||||
Db 144 AspPheCysProGlyIleArgCysTrpValThrGlyTyrPheGlyTyrThrArgGlyGlyGly 163
QY 430 CGCCTCCACCGCCATTTCTCTGAAGCAGTGAAGGTCCCATATGGAAGAACACCAT 489
    ||||| ||||| ||||| ||||| ||||| |||||
Db 164 ProLeuProProTyrSerLeuArgGlyValIleValSerValAlaAspThrGlyThr 183
QY 490 TGTGACGCAAAATACACCTTGGCGCTTACACGCGAGACGACGCTCGCATGCTCGGTGAC 549
    ||||| ||||| ||||| ||||| ||||| |||||
Db 184 CysArgArgAsp-----TyrProGlyProGlyGlySerIleLeuGlnPro 198
QY 550 GACATGCTGTGTGCGCGGAGAACCGCGAGGAGCATGCGCAGGAGGACCTCGGAGGCGCC 609
    ||||| ||||| ||||| ||||| ||||| |||||
Db 199 AspMetLeuCysAlaIleArgGlyPro---GlyAspAlaCysGlnAspAspSerGlyGlyPro 217
QY 610 CTGATGTCAGAGTGAATGACACCTGCTGACAGCGCGGTGCTGACGTGCGGCGGAGGCG 669
    ||||| ||||| ||||| ||||| ||||| |||||
Db 218 LeuValCysGlnValAsnGlyAlaIleTrpValGlnIleGlyThrValSerTrpGlyGlyGly 237
QY 670 TGT 672
    ||||| ||||| ||||| |||||
Db 238 Cys 238

```

Search completed: January 31, 2003, 07:16:39
 Job time : 45 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OW nucleic - protein search, using frame_plus_n2p model

Run on: January 31, 2003, 06:51:34 ; Search time 25.5 Seconds
(without alignments)
5813.305 Million cell updates/sec

Title: US-09-598-982-20
Perfect score: 1458
Sequence: 1 gggccctcgcgaagaagaat.....cgtgaagcggcgcgcgcgtcgt 771

Scoring table: BIOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p-model -DEV=xlh
-Q=/cgn2_1/USPNO.spool/US09598982/unat_27012003_073717_19581/app-query.fasta_1.967
-DB=PIR-73 -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NCR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09598982 -CGN 1.1.15 -Grunat_27012003_073717_19581 -NCPU=6 -ICPU=3
-NO_XLPRY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN_TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: PIR:73:*
2: PIR:1:*
3: PIR:3:*
4: PIR:4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Query Match	Length	ID	Description
1	1368	93.8	275	2	B35863
2	1363	93.5	275	2	tryptase (EC 3.4.2
3	1339	91.8	275	2	tryptase (EC 3.4.2
4	1232.5	84.5	274	2	tryptase (EC 3.4.2
5	1092	74.9	276	2	tryptase (EC 3.4.2
6	1077	73.9	275	2	tryptase (EC 3.4.2
7	1075	73.7	270	2	tryptase (EC 3.4.2
8	1057	72.5	273	2	tryptase (EC 3.4.2
9	1055	72.4	274	2	tryptase (EC 3.4.2
10	1005	68.9	237	2	tryptase (EC 3.4.2
11	808	55.4	230	2	tryptase (EC 3.4.2
12	684.5	46.9	269	2	tryptase (EC 3.4.2
13	512	35.1	343	1	tryptase (EC 3.4.2
14	467	32.0	625	1	tryptase (EC 3.4.2

15	451	30.9	638	1	KOHUP	plasma kallikrein
16	448	30.7	455	2	A61545	plasma (EC 3.4.21
17	447	30.7	638	1	KOMSP	plasma kallikrein
18	446.5	30.6	812	1	PLBO	plasma (EC 3.4.21
19	445	30.5	638	1	KORTPL	plasma kallikrein
20	444.5	30.5	271	1	ELRT2	pancreatic elastas
21	443	30.4	245	1	KYBOB	chymotrypsin (EC 3
22	441	30.2	460	2	B61545	plasma (EC 3.4.21
23	439.5	30.1	269	2	B26823	pancreatic elastas
24	437	30.0	1524	2	T30337	pancreatic elastas
25	430.5	29.5	271	2	A25528	polyprotein - Altr
26	429	29.4	810	2	PLHU	pancreatic elastas
27	428	29.4	810	2	B30848	plasma (EC 3.4.21
28	427	29.3	263	2	A21195	chymotrypsin (EC 3
29	427	29.3	266	1	ELRG	pancreatic elastas
30	426	29.2	263	2	A31299	pancreatic elastas
31	422	28.9	812	1	KYRTB	chymotrypsin (EC 3
32	420.5	28.8	812	1	PLMS	plasma (EC 3.4.21
33	420	28.8	786	1	A47547	serine proteinase
34	419	28.7	245	1	KYBOA	chymotrypsin (EC 3
35	419	28.7	558	2	UC5878	plasma hyaluronan-
36	419	28.7	790	1	PLPG	plasma (EC 3.4.21
37	416.5	28.6	246	2	B25528	trypsin (EC 3.4.21
38	414.5	28.4	246	1	TRRT2	trypsin (EC 3.4.21
39	414.5	28.4	264	2	TRRT2	trypsin (EC 3.4.21
40	413.5	28.4	269	2	C26823	chymotrypsin-like
41	413.5	28.4	367	2	JE0104	pancreatic elastas
42	412.5	28.3	246	1	TRRT1	testicular serine
43	412	28.3	258	2	S70439	trypsin (EC 3.4.21
44	412	28.3	267	4	A56615	pancreatic elastas
45	411	28.2	1420	2	A32869	probable pancreati
						apolipoprotein(a)

ALIGNMENTS

RESULT 1
B35863
tryptase (EC 3.4.21.59) II precursor - human
N:Alternate names: tryptase beta
C:Species: Homo sapiens (man)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C:Accession: B35863; A37193; I59473
R:Vanderlisse, P.; Ballinger, S.M.; Tam, E.K.; Goldstein, S.M.; Craik, C.S.; Caughey,
Proc. Natl. Acad. Sci. U.S.A. 87, 3811-3815, 1990
A:Title: Human mast cell tryptase: multiple cDNAs and genes reveal a multigene serine
proc. Natl. Acad. Sci. U.S.A. 87, 3811-3815, 1990
A:Reference number: A35863; MUID:90251647; PMID:2187193
A:Accession: B35863
A:Molecule type: mRNA; DNA
A:Residues: 1-275 <VAN>
A:Cross-references: GB:M33492; NID:g3339982; PIDN:AAA36779.1; PID:g3339983
A>Note: Residues 2-275 are derived from mRNA; residue one was inferred from the genom
R:Miller, J.S.; Moxley, G.; Schwartz, L.B.
J. Clin. Invest. 86, 864-870, 1990
A:Title: Cloning and characterization of a second complementary DNA for human tryptas
A:Reference number: A37193; MUID:90369005; PMID:2203827
A:Accession: A37193
A:Molecule type: mRNA
A:Residues: 1-275 <MIT>
A:Cross-references: GB:M37488; NID:g179583; PIDN:AAA51843.1; PID:g179584
R:Blom, T.; Hellman, L.
Scand. J. Immunol. 37, 203-208, 1993
A:Title: Characterization of a tryptase mRNA expressed in the human basophil cell lin
A:Reference number: I59473; MUID:93166209; PMID:8434231
A:Accession: I59473
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-275 <RES>
A:Cross-references: GB:S5551; NID:g265666; PIDN:AAD13876.1; PID:g4261576
A:Experimental source: basophil cell line KU812
C:Genetics:
A:Gene: GDB:TPS1
A:Cross-references: GDB:125890; OMIM:191080
A:Map position: 16pter-16qter

Db	131	ValAsnValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro	150
QY	379	CCGGGAGTCCGCTCGGGTCACTGGCTGGGGCGATGGACAATGATGAGCGCTTCCA	438
Db	151	ProGlyMetProCysTrpValThrGlyThrProLysPheValAspAsnAspGluArgLeuPro	170
QY	439	CCGGCATTTCTCTGTAGACACAGGTGAAGGTCCCATATGTGAAACACATTTGTGACSCA	498
Db	171	ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla	190
QY	499	AAATACACCTTTGGCGGCTTACACGGGAGACAGACGTCGCCATCTGTCGAGACATCTGTG	558
Db	191	LysTyrHisIleuGlyAlaTyrThrClyAspAspValArgIleValArgAspAspMetLeu	210
QY	559	TGTGCGGGAGACACCCGGAGGAGATCATGCCAGGCGGACCTCGGAGGGCGCTGGTGTGC	618
Db	211	CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValLys	230
QY	619	AAGGTGAATGGACACCTGGCTCAGGCGGGCGCTGTGTACACTGGGCGAGCGCTGTGCCAG	678
Db	231	LysValAsnGlyThrThrProLeuGlnAlaGlyValAlaValSerTrpGlyGlnGlyCysAlaGln	250
QY	679	CCCAACCGGCGCTGCATCTTACACCCGTGTCACTACTTGGACTGGATCCACACACTAT	738
Db	251	ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr	270
QY	739	GTCCCCAANAACCGC	753
Db	271	ValProLysLysPro	275

OY	19	ATCCGCGGGGCGACAGAGGGCCCCCGAGAGCAAGTGGCCCTGGACAGTACGCTAGAGATC	78
Db	31	llevaIlgylgIglblalaIarIarSerIystrPrroTgblvalIserIleuIarval	50
OY	79	CACGGCCCATCTAGTGGATCGACCTTGTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGTG	138
Db	51	ArgaArPaIyglYtrPmetHISphecysglYglYserIleuIleHISproGlnItrPvalIleu	70
OY	139	ACCGCGCGCGGCGTGGGAGCCGGACGCTCAAGAGTGTGGCCGCCCTCAGGGTGGCACTG	198
Db	71	ThraIaalaIHIScYvalaIglYproAsPvalIySaSPleuIaIalaIeuaIyglIleu	90
OY	199	CGGAGCGACACCTTACTTACCAGGACCAAGCTGCTGCCGGGTACAGAGATCATCTGTGAC	258
Db	91	ArgIgluIglIHISleuYtrYtIglInaSPleuIleuProlSerArgIIleIleValHIS	110
OY	259	CCACAGTTCACACCGCCCGAGATCGGAGCGGACATCGCCCTGCTGAGACTGAGAGGCGG	318
Db	111	ProGlnPheYtrThIaIaGlInIleGlYalaSPleIalaIeuleuIleuGlInIuIPro	130
OY	319	GTGAAGTCTCCAGCCACGCTCCACACGGTCAACCTGCCCCCTGCCCTCAGAGACCTTCCCC	378
Db	131	ValaInValIserSerHISvalIHISThrValThIreuProlaIaSerGIuThrPhePro	150
OY	379	CCGGGAGATGCCGTGTGGGTCACTGTGCTGGGGGCGATGTGGACAATGATGAGCGCTCCCA	438
Db	151	ProIglYmeIProCYstrIYvalThrlYtrPglYasPvaIaSPaIaSPaIaIargIleuPro	170
OY	439	CGGCATTTCCTCTAAACACAGGTGAAGGTCCCGCAATATGAGAAAACCACTTTGTATACGA	498

RESULT 3
 C35863
 trypsinase (EC 3.4.21.59) III precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 03-Feb-1994 #sequence-revision 03-Feb-1994 #text-change 15-Jun-2001
 C:Accession: C35863; A35863; A38893
 R:VanderSllice, P.; Ballinger, S.M.; Tam, E.K.; Goldstein, S.M.; Craik, C.S.; Caughey, G.
 Proc. Natl. Acad. Sci. U.S.A. 87, 3811-3815, 1990
 A:Title: Human mast cell trypsinase: multiple cDNAs and genes reveal a multigene serine p
 A:Reference number: A35863; MUID:90251647; PMID:2187193
 A:Accession: C35863
 A:Molecule type: mRNA
 A:Residues: 9-275 <VAN>
 A:Accession: A35863
 A:Molecule type: DNA
 A:Residues: 1-9 <VA>
 A:Cross-references: GB:M33494; NID:g9327804; PIDN:AAC83172.1; PID:g339977
 A:Note: the first nine residues of this sequence are inferred from genomic DNA of trypt
 R:VanderSllice, P. submitted to GenBank, April 1990
 A:Reference number: A38893
 A:Accession: A38893
 A:Molecule type: mRNA
 A:Residues: 9-131, 'K', 132-275 <VA3>
 A:Cross-references: GB:M33493; NID:g339984; PIDN:AAA36780.1; PID:g339985
 C:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; serine proteinase; zymogen
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-30/Domain: activation peptide #status predicted <ACT>
 F:31-275/Product: trypsin I #status predicted <MAT>
 F:31-261/Domain: trypsin homology <TRI>
 F:74,121,224/Active site: His, Asp, Ser #status predicted

RESULT 4
A45754
tryptase (EC 3.4.21.59) alpha precursor - human
CtSpecies: Homo sapiens (man)
CjDate: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 08-Sep-1997
CjAccession: A45754; B37193
RjMiller: J.S.; Westlin, E.H.; Schwartz, L.B.
UjClin. Invest. 84, 1188-1195, 1989
AjTitle: Cloning and characterization of complementary DNA for human tryptase.
AjReference number: A45754; MUID:90009311; PMID:2677049

Alignment Scores:	
Pred. No.:	1 09e-84
Score:	1339.00
Percent Similarity:	97.96%
Best Local Similarity:	97.96%
Query Match:	91.84%
DB:	2
Length:	275
Matches:	240
Conservative:	0
Mismatches:	5
Indels:	0
Gaps:	0

A:Molecule type: mRNA
A:Residues: 1-274 <MIL>
A:Cross-references: GB:M30038
R:Miller, J.S.; Moxley, G.; Schwartz, L.B.
J. Clin. Invest. 86, 864-870, 1990
A:Title: Cloning and characterization of a second complementary DNA for human tryptase
A:Reference number: A37193; MUID:90369005; PMID:2203827

US-09-598-982-20 (1-771) x C35863 (1-275)

A; molecule type: mRNA

A:Residues: 1-274 <M12>
A:Cross-references: GB:M30038
A:Note: the sequence from Fig. 4 is inconsistent with that from Fig. 2 in having 205-Pro
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase; zymogen
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-30/Domain: activation peptide #status predicted <ACT>
F:31-274/Product: trypsin I #status predicted <MAT>
F:74,120,223/Active site: His, Asp, Ser #status predicted

Alignment Scores:

Prod. NO.:	2,34e-77	Length:	274
Score:	1232.50	Matches:	220
Percent Similarity:	92.65%	Conservative:	7
Best Local Similarity:	89.80%	Mismatch:	17
Query Match:	84.53%	Indels:	1
DB:	2	Gaps:	1

US-09-598-982-20 (1-771) x A45754 (1-274)

QY	19	ATCTCGGGGGGTCAAGAGAGCCCCAGAGCAAGTGGCCCTGGACAGTACGCTGAGATC	78
Db	31	ILEVALIGLIGLGNGLALAPRALSERLSTPRPTGRGLNALSERLEUARGVAL	50
QY	79	CACGGCCCAATCGATGATCACTTCGGGGGCTCCATCCACCCCAAGGGGTGGT	138
Db	51	ARGAPRALGTGTTRPMENHISPHECYSELGLYSERLEULENHSPTGGLNTPRALLEU	70
QY	139	ACCGCCGGCGGTGCGGAGCCGAGCCGTCAAGATGTGGCCGCTCAGGATGCACTG	198
Db	71	THRALAALNHSYLSLEUGLYPTOASRYALLYSASRLAULATHRLEUGVALASN---	89
QY	159	CGGAGACAGACCTTATACAGAGACCAAGTGTGCGGGGTACAGAGATATGCTGCAC	258
Db	90	SERGLYTHRHSLEUTYTYTGLNAPRLNLEUPROVALSERARGILEMETVALHIS	109
QY	259	CCACAGTCTACACCGCCCGCATCGAGCCGAGCATCGCTGCTGAGCTGGAGAGCCG	318
Db	110	PROGLNPHETYLLEILEGLNTHRGLYALAPRIEALALEULENGLNLEUGLNSUR	129
QY	319	GTGAAGTCTCCAGCCAGCCAGTCCACAGGTTCAACCTGGCCCTGCTGCAAGACCTTCCC	378
Db	130	VALAENILIESERFERGVALHISTHRVALMETLEUPRODALASERGLUTHRPHEPRO	149
QY	379	CGGGGATGCGGTGGGTGCTGAGTGGGGGCGATGTGCAADGATGAGCCGCTCCCA	438
Db	150	PROGLYMERPROCYSTRVALTHRGLYTRPGLYASPRVALASPRASNPGLNPROLEUPRO	169
QY	439	CGGCATTCTCTGTAACACAGGTGAAGTCCCAATAAAGGAAACACATTTGTGACGCA	498
Db	170	PROPRORPHEROLENLYSGNLVALYVALPROILEMETGLNASHNHSILECYASPRALA	189
QY	499	AAATACCACTTTGGCGCTTACACGGGAGACGACGTCCCATGTGCTGGACACATGCTG	558
Db	190	LYSTYRHSLSLEUGLYALATYTRGLYASPRVALARGILEILEATGASPRMETLEU	209
QY	559	TGTGCGCGGGAACACCGGAGGACATCAGGCCAGGGGCGCATCGGAGGGCCCTGGTGGC	618
Db	210	CYSALAIGLYASHSERGLNATGASPRERCYSLSGLYASPRSEGLYGLYPRILEUVALCYS	229
QY	619	AAGGTGAATGGACACTGGCTGCAGGCGGGCGGTGATGACACTGGGAGAGGCTGTGCCAG	678
Db	230	LYSVALANGLYTHRTYRPLEUGNLALGLYVALVALISERTPRASRGLNGLCYSLALDELN	249
QY	679	CCCAACCGGCTGGCATGTACACCCGTGTCACTTACTTGGACTGGATGCCACCATAT	738
Db	250	PROASNATGPROGLYILETGTYTHRARGVALTHRTYTYLLEUASPTRILENHSISTYR	269
QY	739	GTCCCCAAAAAGCGG 753	
Db	270	VALPROLYSLYSPRO 274	

```

RESULT 5
A38654
    mast cell proteinase 6 (EC 3.4.21.-) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 21-Feb-1992 #sequence_revision 17-Feb-1994 #text_change 22-Jun-1999
C:Accession: A38654; B38654; D35646; I59478
R:Reynolds, D.S.; Guiley, D.S.; Austen, K.F.; Serafin, W.E.
J Biol. Chem. 266: 3847-3853, 1991
A>Title: Cloning of the cDNA and gene of mouse mast cell protease-6. Transcription by
A:Reference number: A38654; MUID:91139682; PMID:1995638
A:Accession: A38654
A:Molecule type: DNA
A:Residues: 1-276 <RE>
A:Cross-references: GB:M57625; NID:g200506; PIDN:AAA39987.1; PID:g200507
A>Note: the authors translated the codon CGC for residue 24 as Ala, GAG for residue 3
s Gly, GAG for residue 148 as Gly, GAG for residue 168 as Gly, and GAA for 185 as Gly
A:Accession: B38654
A:Molecule type: mRNA
A:Residues: 1-276 <REZ>
A:Cross-references: GB:M57626; NID:g200508; PIDN:AAA39988.1; PID:g200509
R:Reynolds, D.S.; Stevens, R.L.; Lane, W.S.; Carr, M.H.; Austen, K.F.; Serafin, W.E.
Proc. Natl. Acad. Sci. U.S.A. 87, 3230-3234, 1990
A>Title: Different mouse mast cell populations express various combinations of at least two distinct serine proteinases
A:Reference number: A35646; MUID:90222202; PMID:2336280
A:Accession: D35646
A:Molecule type: protein
A:Residues: 32-54 <RE3>
R:Huang, R.; Abirink, M.; Gohl, A.E.; Nilsson, G.; Aveskog, M.; Larsson, L.G.; Nilsson
Scand. J. Immunol. 38, 359-367, 1993
A>Title: Expression of a mast cell tryptase in the human monocytic cell lines U-937 a
A:Reference number: I59478; MUID:94023807; PMID:8210998
A:Accession: I59478
A>Status: preliminary: translated from GB/EMBL/DDBT
A:Molecule type: mRNA
A:Residues: 1-276 <RES>
A:Cross-references: GB:L31853; NID:g473480; PIDN:AAA39725.1; PID:g473481
C:Genetics:
A:Gene: MCP-6
A:Introns: 24/1: 79/2: 168/1: 222/3
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase; zymogen
F:1-21/Domains: signal sequence #status predicted <SIG>
F:22-31/Domains: activation peptide #status predicted <ACT>
F:32-276/Product: mast cell proteinase 6 #status experimental <MAT>
F:32-268/Domains: trypsin homology <TRY>
F:75,122,225/Active site: His, Asp, Ser #status predicted

Alignment Scores:
Pred. No.:      1,11e-67      Length:      276
Score:          1092.00       Matches:     192
Percent Similarity: 84.80%   Conservative: 20
Best Local Similarity: 76.80% Mismatches:    36
Query Match:    74.90%      Indels:      2
DB:             2           Gaps:        1

US-09-598-982-20 (1-771) x A38654 (1-276)

QY      4 CCCCTCGAGAAAGA-----ATGCTGGGGGTACGAGGCCCCCAAGACAAGTGGCCC 57
         ||| .....||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      25 ProlaanaanglnatValgIylleValdllylglnstgualdaSerlnuSerLystrPro 44
         ||| .....||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      58 TGGCAGTGACCTAGAGAGTCACGCGCCCATPGAGATGACTCTGCGGGGCGTCCCTC 117
         ||| .....||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      45 TrpglnvalSerleuAdnrgPheUlysleuAsntYtrprlleHisPhcysglglySerleu 64
         ||| .....||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      118 ATTCACCCCCCAGTGGGTCTACGCGCGCGGTGCCTGGGACCGAGTCACGAATCTG 177
         ||| .....||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      65 lHeHsPrOGIntPrvalleutThrAlalaHlsicysvalGIyPrOHStlleYserPro 84
         ||| .....||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      178 GCCGCCCTCAGGGTGCACACTGCGGGAGACAGCACTTAATAACAGGACAGCTCTGCCG 237
         ||| .....||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      85 GluleupheaayValGlInleuAdnglnelInleuYtyrYrGLyaspGlnleuSer 104
         ||| .....||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


Qy	586	TGCGAGGGCCACATCCGAGAGGGCCCCCTGCTGTGCAAGCTAAATGGACACCTGGCTGCAGGCG	645
Qy	182	CysGlnGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAla	201
Qy	646	GGCGTGTGACGCTGGGGGCGAGGGCTGTGGCCGACGCCCAACGGCGCTGGACATTCACACCGGT	705
Db	202	GlyValValSerTrpGlyAspGlyCysAlaLysProAsnTrpGlnGlyLeuThrArg	221
Qy	706	GTCACCTACTACTGTGACTGTGATCCACCACTATGTCTCCCAAAAGCG	753
Db	222	ValThrSerTrpLeuAspTrpIleHisGlnTrpValProGlnGlyPro	237
RESULT 11			
I48685			
mast cell proteinase 6 (EC 3.4.21.-) precursor - mouse			
C:Species:	Mus musculus (house mouse)		
C:Date:	02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 22-Jun-1995		
C:Accession:	I48685; S43172		
R:Huang, R.; Hellman, L.			
Immunogenetics	40, 397-414, 1994		
A:Title:	Genes for mast-cell serine proteinase and their molecular evolution.		
A:Reference number:	I48684, MID:J95048582, PMID:7959952		
A:Accession:	I48685		
A:Status:	preliminary; translated from GB/EMBL/DBJ		
A:Molecule type:	mRNA		
A:Residues:	1-230 <RES>		
A:Cross-references:	EMBL:X78542; NID:g468809; PIDN:CA55288.1; PID:g468810		
C:Superfamily:	trypsin; trypsin homology		
C:Keywords:	hydrolase; serine proteinase		
C:33-230/Domain:	trypsin homology #status atypical <TRY>		

Alignment Scores:			
Pred. No.:	4.03e-48	Length:	230
Score:	808.00	Matches:	144
Percent Similarity:	81.82%	Conservative:	18
Best local Similarity:	72.73%	Mismatches:	34
Query Match:	55.42%	Indels:	2
DB:	2	Gaps:	1
US-09-598-982-20 (1-771) x I48685 (1-230)			
QY	4	CCCCTCGAGAAAAGA-----ATCGTGGGGGTGCGAGGCCCCCAGGACCAAGTGGCCC	57
Db	25	Prokalsnnglnatgvalgyllevalgcylihsctlnalsersgluserlystppro	44
QY	58	TGCGAGGTGAGCCGTGAAGAGTCCACGGCCCACTGATGCATTCTGGGGGCGCTCC	117
Db	45	TrpInValSerLeuAtgRhpheLyLeuAsnTyrTrpIleHisPheCysGlyLySerLeu	64
QY	118	ATTCACCCCCAGTGGGTGCTGACCGCGCGCGCTGCTGGAGCCGACGTCAGAGATCTG	177
Db	65	IleHisProGlnTrpValLeuThrAlaAlaHisCysValdYpRhoHisIleLysSerPro	84
QY	178	GCCGCCCTGAGGGTGCAGTGCAGGAGCAGACCTACTACGACGAGACAGCTGCTCCG	237
Db	85	GlnLeuRheaRgValGlnLeuAtgGlnIntYrLeuTyrTyrGlyAspGlnLeuSer	104
QY	238	GTCGACAGGATCATGTGCACGCCACAGTCTACACGCCGCCAGATCGAGCGGACATGCC	297
Db	105	LeuAsnArgIleValValHisProHisIleTyrThrAlaGlnGlyAlaAspValAla	124
QY	298	CTGTGTGAGCGTGGAGGAGCCGGTGAAGTGTCCAGGCCACGTCACACGGGTACCCCTGCC	357
Db	125	LeuLeuGlnLeuGlnValProValAsnValSerThrHisIleHisProlIleSerLeuPro	144
QY	358	CCGTGCTCAGAGACCTTCCGCCCGGAGATGCCGCTGGGTCACTGGCTGGGCGATGTG	417
Db	145	ProkalsersglutrnRhpheProTrogIYrHisCysTrrpValThrGlyTrpGlyAspIle	164
QY	418	GACAAATGATGAGCGCTCCACCGGCATTTCTCTGGAAGCAGGTGAAGGTCCCATATAG	477
Db	165	AspAsnAspGlnProlLeuProGProTyrProLeuLysGlnValLyValProlIleVal	184

```

QY 478 GAAACCCACATTTTGAGACGCCAAATACACCACTTGGCGGCTTACAGGGAGACAGACTCCGC 537
      ||||| ..... ||||| ||| ||||| |||||
Db 185 GUAASerLeuCySAsPArglyStryrHISrThrglyLeuTYrThnGlyAsPAsPhePro 204
      ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 538 ATCGTCCTGACGACATCTGTCGCCGGGAACACCCGGAGGAGACTATGCCAG 591
      ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 205 ILEVALHAsPglyMetLeuCySAlaglySntHrArgAsPserCysGln 222
      ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
B32410
mastocytoma proteinase (EC 3.4.21.-) precursor - db9
C:Species: Canis lupus familiaris (dog)
C:Date: 12-Oct-1989 #sequence_revision 31-Dec-1993 #text_change 10-Sep-1997
C:Accession: B32410
R:Vanderslize, P.; Craik, C.S.; Madel, J.A.; Caughey, G.H.
Biochemistry 28, 4148-4155, 1989
A:Title: Molecular cloning of dog mast cell tryptase and a related protease: structur
A:Reference number: A32410; MUID:89352460; PMID:2504277
A:Accession: B32410
A:Molecule type: mRNA
A:Residues: 1-269 <VAN>
A:Cross-references: GB:M24665; NID:g163984; PID:g163985; GB:J02862
A:Note: the authors translated the codon ACG for residue 114 as Ser
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:1-19/domain: signal sequence #status predicted <SRC>
F:20-269/Product: mastocytoma proteinase #status predicted <MAT>
F:20-259/Domain: trypsin homology <TRY>
F:66,116,217/Active site: His, Asp, Ser #status predicted

```

Alignment Scores:			
Pred. No.:	1.28e-39	Length:	269
Score:	684.50	Matches:	132
Percent Similarity:	65.34%	Conservative:	32
Best Local Similarity:	52.59%	Mismatches:	78
Query Match:	46.95%	Indels:	9
DB:	2	Gaps:	4
US-09-598-982-20 (1-771) x B32410 (1-269)			
QY	19	ATGTGGGGGGGTCAAGAGAGGSSCCAGAGAGCAAGTGGSSCTTGACAGTGAAGTGC	78
		:::	:::
Db	20	evAl g u g y s y l v a l p ro l a n g u l p ro t g n l a s e r l e u d a l g p he	39
QY	79	CACGGC-----CCATGATGATGACACTTCGCGGGGCTCCGTCATCCACCCAG	129
Db	40	s t g e t g y s e r g l u l t p r o l n l s l e c y s g l y s e r l e u l l e n s p ro g l n	59
QY	130	TGGGTGTGACGGCGCGCGCGCTGGGTGGAGCCGAGCAGTCAAGATGTCGCGCCCTCAGG	189
Db	60	p v a l g u l p r o l a l a n s c y u v a l d u e u g l u e n g l u a l l a l l e u d a l g	79
QY	190	GTGCACTGCGGGAGAGCAAGCACTTACTACAGAGACCAAGCTGCTGGCGGTACAGAGATC	249
		:::	
Db	80	v a l g u l g u l l e u d a l g l e u l t y a s p n s a s p n l e u s a s n v a l t p h l u l e	99
QY	250	ATGCGGACCCCAAGATTGTACACCGGCCCAAGATCGGA-----GGGAGATGGCCCTG	300
Db	100	l e a n g l s p r o a n p h a s m e t s e r t p r y l t y t p a s p n l a a s p l e a l e u	119
QY	301	CTGGAAGCTGGAGAGACCGGGTGAAGGTCTCCAGCCACGTCACACGTCACCCCTCCCTC	360
Db	120	e u l s l e u l a l p r o l e u t p h l e u s e r l u s p a l s n v a l s e r l e u l p r o s e r	139
QY	361	GGCTCAGAGACCTTCCCGCGGGGATGGCGCGTGGGTCACCTGGCTGGGGGATGTGGAC	420
Db	140	p r o s e r l e u l l e v a l p r o t g l y e t l e u c y s t p v a l t p h g l y t p c l y a s p l e a l	155
QY	421	AATGATGAAGCCCTCCACACCCCATTTCTCTGTGAAGAGGTGAAGTCCCATATAGAA	480
		:::	
Db	160	a s p n s t p h p r o l e u p r o t g l y n s l e u n g l u a l g u a l p r o l e v a l g y	179

R:McMullen, B.A.; Fujikawa, K.; Davie, E.W.
 Biochemistry 30, 2056-2060, 1991
 A:Title: Location of the disulfide bonds in human coagulation factor XI: the presence of
 A:Reference number: A37940; MUID:91152017; PMID:1998667
 A:Accession: A37940
 A:Molecule type: protein
 A:Residues: 28-33;35-49,'X',51-55,'X',57-63;70-75,'X',77-79;107-109,'X',111-112;132-139;
 ;280-282,'X',284;285-297;313-316,'X',318-319;320-326;'X',328-330;'X',347-349;373,'X',375
 C:Comment: The proenzyme consists of two identical chains linked by one or more disulfide
 he active site, and a heavy chain, which associates with high molecular weight (HMW) kit
 C:Genetics:
 A:Gene: GDB:F11
 A:Cross-references: GDB:119891; OMIM:264900
 A:Map position: 4q35-4q35
 A:Introns: 19/1; 73/2; 109/1; 162/2; 199/1; 252/2; 289/1; 343/2; 379/1; 435/2; 494/1; 52
 C:Function:
 A:Description: catalyzes the proteolytic activation of coagulation factor IX
 A:Pathway: blood coagulation intrinsic pathway
 C:Superfamily: coagulation factor XI; trypsin homology
 C:Keywords: blood coagulation; duplication; glycoprotein; hemophilia C; homodimer; hydro
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-187/Product: coagulation factor Xla heavy chain #status experimental <HCH>
 F:19-108/Domain: apple repeat <AP1>
 F:109-198/Domain: apple repeat <AP2>
 F:199-288/Domain: apple repeat <AP3>
 F:290-379/Domain: apple repeat <AP4>
 F:388-625/Product: coagulation factor Xla light chain #status experimental <LCH>
 F:20-103,514-581,571-599/Disulfide bonds: #status predicted
 F:388-618/Domain: trypsin homology <TRY>
 F:29-Disulfide bonds: interchain #status experimental
 F:46-76,50-56,110-193,136-153,140-146,200-283,226-255,230-236,291-374,317-346,321-327,38
 F:90,126,353,450/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:339/Disulfide bonds: interchain #status predicted
 F:387-388/Cleavage site: Arg-Ile (coagulation factor Xla) #status experimental
 F:431,480,575/Active site: His, Asp, Ser #status predicted
 F:491/Binding site: carboxylate (Asn) (covalent) #status experimental

Alignment Scores:

Pred. No.:	1,18e-24	Length:	625
Score:	467.00	Matches:	99
Percent Similarity:	51.59%	Conservative:	31
Best Local Similarity:	39.29%	Mismatches:	94
Query Match:	32.03%	Indels:	28
DB:	1	Gaps:	6

US-09-598-982-20 (1-771) x KFHU (1-625)

```

OY      7 CTCGAGAAAGATGTGGGGGTAGAGAGCCCGAGAGCACTGGCCGTGAGGTG 66
          ::::: ||||| ||||| ||||| ::::: ||| ::||| ||||| |||||
Db      384 lletysProAtgRlletValglYlThrlAserValAtrglglYlUtrPrProtrpRlnVal 403
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY      67 AGCCGAGAGTCCAGGCGCATACGTGACCTTGGGGGGGCTCCCTCATCCACCC 126
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      404 ThrlsuhHlstrThrSerProthrlnArghlslLeucysglYlSerlletllelYasn 423
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY      127 CAGTGGTGTCCACCGCCGCGCTGCGGAGCCGAGCAAGTCAAGATCTGGCCCTTC 186
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      424 GlntrPrleleuThrlAlAlAlHisCysPhe--TyrGlYvalGlulserProlyslleu 442
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY      187 AGGTGCA-----CTGGGAGAGAGCACTCTACTACAGAGACAGCTGCTGCCG 237
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      443 ArYvalTYrSerGlYlletleuasnGlnserglullelYsgluAsPthSerPheheglY 462
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY      238 GTACAGAGATCATGTGATCCACCATGTTATACAGCCAGATGAGGAGCATGCGC 297
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      463 ValglnglullelletleHisaspelnTrYrYsMetAlaGlulserGlYTrasPillella 482
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY      298 CTGCTGAGCTGAGAGACCGGTGAAGTCTCACACCATCCACGCGTCCAGCTGCC 357
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      483 leuLeuYslleuGlUtrThrValasnYrThraserGlInArGProilecYslleuro 502
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY      358 CCGCTGCTGAGCACTTCCCGCCCGGAGATGCCGCTGCTGGGTCACTGGCGGC----- 411
          ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 503 SerlysglyAspArGAsnVallelYrThrasPcYstrpValThnglyTrpGlYTrArG 522
 OY 412 -----GATGGAGACATGATGAGCGCTCCACCGCATTTCCCTGAACAG 459
 Db 523 LysleuArGAspYslleGlAsn-----ThrlleuGlnlys 534
 OY 460 GTGAAGTCCCATTAAGGAACACATTTGTGACGCCAAATACACACTTGGCGCTTAC 519
 Db 535 AlalYslleProleuValThrasnGlnlucYsglnYsArGTYrArGglYnHlslYsle 554
 OY 520 ACGGAGACAGACGTCCGATGCTCCGACACATGCTGTGGCGGAGAACACCCG--- 576
 Db 555 Thr-----HislyMetlleCysAlaGlYTrArGglngly 566
 OY 577 --AGGACTATGACGAGGAGGACATCCGAGGCCCTGTGTCAGAGTGAATGACAC 633
 Db 587 GlylyAspAlaCylAspserGlYAspserGlYlProleuSerCylslYnHlAsnGlulVal 586
 OY 634 TGGCTGAGCGCGGTGTGCTGAGCTGGGCGAGGCTGTGCCAGCCACCGGCTTGGC 693
 Db 587 TrpHlslleuValglYlletThrSerTrpGlYlucYsAlaGlInArGgluArGProclY 606
 OY 694 ATCTACACCCCTGCTACCTACTACTGACCTGATC 729
 Db 607 ValYrThrasnValValglYrValAspTrpIle 618

RESULT 15
 KOHUP
 N:Altenate names: Kallikrein (EC 3.4.21.34) precursor - human
 N:Alternate names: Kallikrein; plasma prekallikrein
 C:Species: Homo sapiens (man)
 C:Date: 13-Aug-1986 #sequence,revision 13-Aug-1986 #text,change 18-Jun-1999
 A:Accession: A00921; A37939
 R:Chung, D.W.; Fujikawa, K.; McMullen, B.A.; Davie, E.W.
 Biochemistry 25, 2410-2417, 1986
 A:Title: Human plasma prekallikrein, a zymogen to a serine protease that contains fou
 A:Reference number: A00921; MUID:86243359; PMID:3521732
 A:Accession: A00921
 A:Molecule type: mRNA
 A:Residues: 1-638 <CHU>
 A:Cross-references: GB:M3143; NID:q190262; PIDN:AAA60153.1; PID:q190263
 R:McMullen, B.A.; Fujikawa, K.; Davie, E.W.
 Biochemistry 30, 2050-2056, 1991
 A:Title: Location of the disulfide bonds in human plasma prekallikrein: the presence
 A:Reference number: A37939; MUID:91152016; PMID:1998666
 A:Accession: A37939
 A:Molecule type: protein
 A:Residues: 20-27;40-46,'X',48,'H',50,'X',52-70,'H',75-76,'X',78-80;103-113;131-140;1
 ;260-283,'X',285;287-291,'X',293-295;314-317,'X',319-320;321-324,'X',329-333;334-339,
 525;538-551;562,'X',564-567;573,'X',575-576;578-583,'X',585;592-604 <MCM>
 C:Comment: This protein, synthesized in the liver, circulates as a noncovalent comple
 C:Comment: The zymogen is activated by factor Xlla, which cleaves the molecule into a
 are linked by one or more disulfide bonds.
 C:Comment: The enzyme cleaves Lys-Arg and Arg-Ser bonds. It activates, in a reciproc
 Inogen and may also play a role in the renin-angiotensin system by converting proreni
 C:Genetics:
 A:Gene: GDB:KLK3
 A:Cross-references: GDB:127575; OMIM:229000
 A:Map position: 4q35-4q35
 C:Superfamily: coagulation factor XI; trypsin homology
 C:Keywords: blood coagulation; duplication; fibrinolytic; glycoprotein; hydrolase; in
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-638/Product: plasma kallikrein #status predicted <MAT>
 F:20-390/Domain: plasma kallikrein heavy chain #status predicted <HCH>
 F:20-109/Domain: apple repeat <AP1>
 F:110-199/Domain: apple repeat <AP2>
 F:200-389/Domain: apple repeat <AP3>
 F:291-380/Domain: apple repeat <AP4>
 F:391-621/Domain: plasma kallikrein light chain #status predicted <LCH>
 F:21-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,322-328,
 F:127,308,396,453,494/Binding site: carboxylate (Asn) (covalent) #status experimental
 F:318-347,340-345/Disulfide bonds: #status predicted

F:390-391/Cleaveage site: Arg-Ile (coagulation factor XIIa) #status predicted
F:434,483,578/Active site: His, Asp, Ser #status predicted

Alignment Scores:

Pred. No.:	1.49e-23	Length:	638
Score:	451.00	Matches:	96
Percent Similarity:	54.40%	Conservative:	40
Best Local Similarity:	38.40%	Mismatches:	84
Query Match:	30.93%	Indels:	30
DB:	1	Gaps:	8

US-09-598-982-20 (1-771) x K0HUP (1-638)

```
QY 16 AGAATCTGCGGGGTCAGAGGCCCCAGAGACAGTGACCCCTGACAGTGCAGCTGAGA 75
    |||||
Db 390 ArgIleValGIgLYThrAsnSerSerTrpGLuTrpProTrpGlnValSerLeuGln 409
    |||||

QY 76 GTCCAGGGCCCATACTGATGATGCTTCTGCGGGGCTCCCTCATCCACCCAGTGGGTG 135
    |||
Db 410 ValLysLeuThrAlaGlnArgHisLeuGlySerLeuIleGlyHisGlnTrpVal 429
    |||

QY 136 CTGACCGCGCGCGGCTGGCTG---GGACCGGACGTCAAGATCTGCCCGCCCTCAGGGTG 192
    |||||
Db 430 LeuThrAlaAlaHisCysPheAspGlyLeuProLeuGlnAspVal-----TrpArgIle 447
    |||||

QY 193 CAA-----CTCGGGAGACAGCCTTACTACAGACACAGCTGCTGCCGGTCAGC 243
    |||
Db 448 TyrSerGlyIleLeuAsnLeuSerAspIleThrLysAspThrProPheSerGlnIleLys 467
    |||||

QY 244 AGGATCATGTCGACCCACAGTTCTTACACCGCCAGATCGAGCGACATCGCCCTGCTG 303
    |||||
Db 468 GluIleIleHisGlnAsnTyrLysValSerGlnGlyAsnHisAspIleAlaLeuIle 487
    |||||

QY 304 GAGCTGGAGAGCCGGTGAAGGTCTCAAGCCACGTCCACAGGTGACCCCTGCCCTGCC 363
    |||||
Db 488 LysLeuGlnAlaProLeuAsnTyrThrGlnPheGlnLysProIleCysLeuProSerLys 507
    |||||

QY 364 TCAGAGACCTTCCCGCCGGGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 408
    |||||
Db 508 GlyAspThrSerThrIleTyrThrAsnCysTrpValThrGlyTrpGlyPheSerLysGlu 527
    |||||

QY 409 ---GGCGATGTGACATGATGAGCGCTCCACCGCATTTCTCTGAGAGAGGTGAAG 465
    |||||
Db 528 LysGlyGluIleGlnAsn-----IleLeuGlnLysValAsn 539
    |||||

QY 466 GTCCCATTAATGGAACACACATTGTGACGCAAAATACACCTTGCGCGCTACACGGGA 525
    |||||
Db 540 IleProLeuValThrAsnGluGlnCysGlnLysArgTyr----- 552
    |||||

QY 526 GACGACGTCCGATGCTCGAGACATGCTGCTGCGGGGACACCCCG-----AGG 579
    |||
Db 553 GlnAspTyrLysIleThrGln---ArgMetValCysAlaGlyTyrLysGluGlyLys 571
    |||||

QY 580 GACTCATGCCAGGGGACTCCGAGAGGCCCTGCTGCTGCAAGGTGAATGACACTGCGCTG 639
    |||||
Db 572 AspAlaCysLysGlyAspSerGlyGlyProLeuValCysLysHisAsnGlyMetTrpArg 591
    |||||

QY 640 CAGCGGGCGCTGGTACGTGCGGGGCTGTGCTGCCAAGCCACCGGCTGGCATCTAC 699
    |||||
Db 592 LeuValGlyIleThrSerTrpGlyLysCysAlaArgGlnProGlyValTyr 611
    |||||

QY 700 ACCCGTGCACCTACTACTGAGTGCATC 729
    |||||
Db 612 ThrLysValAlaGluTyrMetAspTrpIle 621
    |||||
```

Search completed: January 31, 2003, 06:59:25
Job time : 33.5 secs

OY	17	GAATTCGTGGGGGGTCAGAGAGGCCCCCGACAGACCAAGTGGCCCTGGCAGAGTGGAGCCCTGAAG	76
Db	66	GCAATGTTGGGGGGTCAGAGAGGCCCCCGACAGACCAAGTGGCCCTGGCAGAGTGGAGCCCTGAAG	125
OY	77	TCACAGGGCCCAATTACGTGATGTCACCTTCTGCGGGGGCTCCCTCATCCACCCTCAGTGGGTGC	136
Db	126	TCCGGACCCGATPACTGGATGTCACCTTCTGCGGGGGGGTCCCTCATCCACCCTCAGTGGGTGC	185
OY	137	TGACCGCCGGCGCGTGGGTGGAGCCGGACGTCAGAGATCTGTGGCCGCCCTCAGGGTGCAC	196
Db	186	TGACCGCGCGCCACTGCTGGGGACCGAGCGTCAGAGATCTGGCCGCCCTCAGGGGTGCAC	245
OY	197	TGCGGAGACAGACACTTACTACTACAGAGACAGCTGCTCCGGTGCAGAGATCATCTGCG	256
Db	246	TGCGGGAGCAGACACTTACTACTACAGAGACAGCTGCTCCGGTGCAGAGATCATCTGCG	305
OY	257	ACCACAGTTCTTACACCGCCCGAGATCGAGCGGACATCGCCCTGCTGGAGCTGAGAGAGC	316
Db	306	ACCACAGTTCTTACACCGCCCGAGATCGAGCGGACATCGCCCTGCTGGAGCTGAGAGAGC	365
OY	317	CGGTGAAGGTCTCCAGCCAGCTCCACAGGTCACCTTCCCTCCCTTGCCTCAGAGACCTTCC	376
Db	366	CGGTGAAGGTCTCCAGCCAGCTCCACAGGTCACCTTCCCTCCCTTGCCTCAGAGACCTTCC	425
OY	377	CCCCGGGATGCGCGTGTGGGTCACTGGCTGGGGGATGTGGACAATGATGAGCCCTCC	436
Db	426	CCCCGGGATGCGCGTGTGGGTCACTGGCTGGGGGATGTGGACAATGATGAGCCCTCC	485
OY	437	CACCGCCATTTCTCTGTAAGCAGGTGAAGGTCCCCATTAATGAAAAACCACTTTGTGACG	496
Db	486	CACCGCCATTTCTCTGTAAGCAGGTGAAGGTCCCCATTAATGAAAAACCACTTTGTGACG	545
OY	497	CAAAATACCACCTTGGGCCCTTACACGGGAGACAGAGTCCCGATCGTCCGTGACGACATGC	556
Db	546	CAAAATACCACCTTGGGCCCTTACACGGGAGACAGAGTCCCGATCGTCCGTGACGACATGC	605
OY	557	TGTGTGCCGGGAACCCCGGAGGAGCATCATGACAGGCGACTCCGAGAGGGCCCTGCTGT	616
Db	606	TGTGTGCCGGGAACCCCGGAGGAGCATCATGACAGGCGACTCCGAGAGGGCCCTGCTGT	665
OY	617	GCAAGTGAATGGCACCTGCGCTGCAGGCGGGCGGTGTCACTGGGGCGAGGGCTGTGCCC	676
Db	666	GCAAGTGAATGGCACCTGCGCTGCAGGCGGGCGGTGTCACTGGGGCGAGGGCTGTGCCC	725
OY	677	AGCCCAACCGCGCTGGCATATACACCGCTGTCACTACTGTGGACGTGATCCACCACT	736
Db	726	AGCCCAACCGCGCTGGCATATACACCGCTGTCACTACTGTGGACGTGATCCACCACT	785
OY	737	ATGTCCCCCAAAAAGCCGTGAAGCGGCGC	763
Db	786	ATGTCCCCCAAAAAGCCGTGAAGCGGCGC	812
RESULT 2			
US-09-954-456-2242/c			
: Sequence 2242, Application US/09954456			
: Patent No. US20020115057A1			
GENERAL INFORMATION:			
: APPLICANT: Young, Paul			
: TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc			
: FILE REFERENCE: 689290-76			
: CURRENT APPLICATION NUMBER: US/09/954, 456			
: CURRENT FILING DATE: 2001-09-18			
: PRIOR APPLICATION NUMBER: US/60/233, 617			
: PRIOR FILING DATE: 2000-09-18			
: PRIOR APPLICATION NUMBER: US/60/234, 052			
: PRIOR FILING DATE: 2000-09-20			
: PRIOR APPLICATION NUMBER: US/60/234, 923			
: PRIOR FILING DATE: 2000-09-25			
: PRIOR APPLICATION NUMBER: US/60/235, 134			
: PRIOR FILING DATE: 2000-09-25			

```

: PRIOR APPLICATION NUMBER: US/60/235,637
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,638
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,711
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,720
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,840
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,863
: PRIOR FILING DATE: 2000-09-27
: NUMBER OF SEQ ID NOS: 2276
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 2242
:
: LENGTH: 616
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:
: NAME/KEY: misc_feature
:
: OTHER INFORMATION: n-a,t,g or c
US-09-954-456-2242

```

Query Match 28.98; Score 223.2; DB 10; Length 616;

Best Local Similarity 93.5%; Pred. No. 8.2e-44;
Matches 286; Conservative 0; Mismatches 15; Indels 5; Gaps 5;

Oy	460	GTAAAGTCCCATTAATAGGAACCAACCATTTGTACGCAAAATTAACCACTTGGCGCTAC	5193
Oy	607	GTGAAGGTCCCATTAATATGG - AAACACACTTTGTATAGCG - AAATTAACCACTTTGGGGCTTAC	5500
Oy	520	ACGGGAGACGACGTCGCCATCGTCCGTG - ACGACATCTGTGTGCCGGGAACACCCGGAG	578
Db	549	ACGGGAGACGACGTCGCCATCATCTCGTGAACGACATCTGTGTGCCGGGAACA - CCGGAG	491
Oy	579	GGACATCATGCCAAGGGGCG - ACTCCGGAGGGGCCCCGTGGTGCAAGGTGAATAGGACACCGCGC	637
Db	430	GGACATCATGCCAAGGGGCGAAGCTCTGTGAGGGCCCCCTGTGTGTGCAAGGTGAATAGGACACCTGGC	4311
Oy	638	TGCAGGCGGGCGTGGTCACTGTGGGGCGAGGGCTGTGGCCACACCCAGCGGCTGTGGCATCT	637
Db	430	TACAGGCGGGGCGTGGTCACTGTGGGANGAGGACGTGTGGCCAGCCCAACCGGCGCTGTGGCATCT	3712
Oy	698	ACACCCGTTGTCACCTACTACTTGGAGCTGGATCCACACACTATGTGTCCCAAAAAGCCGTGAA	757
Db	370	ACACCCGTTGTCACCTACTACTTGGAGCTGGATCCACACACTATGTGTCCCAAAAAGCCGTGGAG	3111
Oy	758	GGGGGCC 763	
Db	310	TCAGGC 305	

```

RESULT 3
US-10-041-006A-1
? Sequence 1, Application US/10041006A
? Patent No. US20020168754A1
? GENERAL INFORMATION:
? APPLICANT: Andrade-Gordon, Patricia
? APPLICANT: Darrow, Andrew
? APPLICANT: Qi, Jian-Shen
? TITLE OF INVENTION: DNA encoding the novel human serine
? TITLE OF INVENTION: protease T
? FILE REFERENCE: ORT-1032
? CURRENT APPLICATION NUMBER: US/10/041,006A
? CURRENT FILING DATE: 2002-01-07
? NUMBER OF SEQ ID NOS: 11
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 1
? LENGTH: 1110
? TYPE: DNA
? ORGANISM: Homo sapiens
US-10-041-006A-1

```

```

Query Match          26.1%; Score 201.4; DB 9; Length 1110;
Best Local Similarity 57.4%; Pred. No. 1.2e-38;
Matches 433; Conservative 0; Mismatches 301; Indels 21; Gaps 3;

QY 12 GAAAGAAATCGTCGGGGGTGACAGAGGCCCCAGAGCAAGTGGCCCTGGCAGTAGCCT 71
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 131 GAACCGAAATGTTGGGGGAGAGACACGAGAGGGCGAGTGGCCCTGGCAAGTCAGCAT 190
QY 72 GAGAGTCACGGCCCATATGATGATGCTTTCGGGGGGCTCCCATCATCACCCAGAG 131
    | 191 -----CGAGCGCACGGAAGCCACTTTCGGGGGAGACTTCATCGCGAGTGTG 241
QY 132 GGTGCTGACCGCGCGCGCTGCGTGGAGCCGAGCTCAAGATCTGGCC-----GCCCT 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 GGTCTGACGGCTGGCGACATGCTCCGCAACACTCTGAGAGTGCTCCCTGTACAGGTCT 301
QY 186 CAGGTGCACTGCGGGAGACAGCAGCTCTACTACAGAGACAGCTGCTCCGCTACGAC 245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 GCTGGGGGCGAAGGCAAGTGTAGTCAGCGCGGAGCACAGCTATGTATGCCGGGTAGGCA 361
QY 246 GATCATCGTCGACCCACACATGTTCTACACCGCCAGATCGGAGGGAGATGGCCCTGTGA 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 GGTGAGAGACAAACCCCTTACCAAGGACAGCGCTTCAGCCTGACGTGGCCCTGTGA 421
QY 306 GCTGAGAGACCGGTGAAGGTCTCCAGCCAGCTCCACAGAGTCCACCTGCCCCCTGCTTC 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 GCTGAGGACCAAGTCCCTTACCAATATACATCTCCCGTGTGCTGCTGACCCCTC 481
QY 366 AGAGACTTCCCCCGGGGATGCCGTGGTCTACTGCTGGGCTGGGGGATGTGACATATA 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 482 GGTGATCTTTGAGACGGGAGTGAATGCTGCTGGGCTGGGGAGGCCCAAGTAGAGA 541
QY 426 TGAGCGCTCCCAACCGCATTTCTCTGAAGCAGGTGAAGTCCCATATATGAAAGCA 485
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 542 AAGCTCTCGCCCGAAGCCGGATCTGCAAGAACTCGCTGTGCCATATGACACAC 601
QY 486 CATTTGTGACGCAAAATACCACTTGGCGCTACACAGGAGAGAGAGCTCCGATGCTCG 545
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 602 CAAGTCAACATGCTCTACAGCAAGACACGAGTTGGCTACCAACCCAAACATCAA 661
QY 546 TGACGACATGCTGTGTGCGG-----GAACACCCGGAAGGACTATGCCAGGGGACTTC 599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 662 GATGACATGCTGTGTGCGGCTTCGAGAGGAGGCAAGAGATGCTCGAAGGGGCACTC 721
QY 600 CGAGGGGCCCCGTGTGTGCAAGTGAATGCACTGGCTGCAAGGGGGGTGTCAGCTG 659
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 722 GGGCGGCCCCCTGTGTGTGCTGTGAGTGTGCTGTGAGGGGGGGGTATATGCTG 781
QY 660 GGGCGAGGGCTGTGCGCACGCCCAACCGGCTGGCATCTACACCGGTGTGCTACTACTT 719
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 782 GGGTGTAGGGCTGTGCGCCGACAGACCGCCAGGTGTCTACATCCGTGTGACCGCCACCA 841
QY 720 GGACTGTGATCCACACTATGTGTCCTCCCAAAAGCCGT 754
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 842 CAAGTGTGATCATCGATCATCCCAAACTGCAGT 876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
US-10-040-655-1
; Sequence 1, Application US/10040655
; Patent No. US20020146805A1
; GENERAL INFORMATION:
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/10/040,655
; CURRENT FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1

```

```

; LENGTH: 1110
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-655-1

Query Match          26.1%; Score 201.4; DB 12; Length 1110;
Best Local Similarity 57.4%; Pred. No. 1.2e-38;
Matches 433; Conservative 0; Mismatches 301; Indels 21; Gaps 3;

QY 12 GAAAGAAATCGTCGGGGGTGACAGAGGCCCCAGAGCAAGTGGCCCTGGCAGTAGCCT 71
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 131 GAACCGAAATGTTGGGGGAGAGACACGAGAGGGCGAGTGGCCCTGGCAAGTCAGCAT 190
QY 72 GAGAGTCACGGCCCATATGATGATGCTTTCGGGGGGCTCCCATCATCACCCAGAG 131
    | 191 -----CGAGCGCACGGAAGCCACTTTCGGGGGAGACTTCATCGCGAGTGTG 241
QY 132 GGTGCTGACCGCGCGCGCTGCGTGGAGCCGAGCTCAAGATCTGGCC-----GCCCT 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 GGTCTGACGGCTGGCGACATGCTCCGCAACACTCTGAGAGTGCTCCCTGTACAGGTCT 301
QY 186 CAGGTGCACTGCGGGAGACAGCAGCTCTACTACAGAGACAGCTGCTCCGCTACGAC 245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 GCTGGGGGCGAAGGCAAGTGTAGTCAGCGCGGAGCACAGCTATGTATGCCGGGTAGGCA 361
QY 246 GATCATCGTCGACCCACACATGTTCTACACCGCCAGATCGGAGGGAGATGGCCCTGTGA 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 GGTGAGAGACAAACCCCTTACCAAGGACAGCGCTTCAGCCTGACGTGGCCCTGTGA 421
QY 306 GCTGAGAGACCGGTGAAGGTCTCCAGCCAGCTCCACAGAGTCCACCTGCCCCCTGCTTC 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 GCTGAGGACCAAGTCCCTTACCAATATACATCTCCCGTGTGCTGCTGACCCCTC 481
QY 366 AGAGACTTCCCCCGGGGATGCCGTGGTCTACTGCTGGGCTGGGGGATGTGACATATA 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 482 GGTGATCTTTGAGACGGGAGTGAATGCTGCTGGGCTGGGGAGGCCCAAGTAGAGA 541
QY 426 TGAGCGCTCCCAACCGCATTTCTCTGAAGCAGGTGAAGTCCCATATATGAAAGCA 485
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 542 AAGCTCTCGCCCGAAGCCGGATCTGCAAGAACTCGCTGTGCCATATGACACAC 601
QY 486 CATTTGTGACGCAAAATACCACTTGGCGCTACACAGGAGAGAGAGCTCCGATGCTCG 545
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 602 CAAGTCAACATGCTCTACAGCAAGACACGAGTTGGCTACCAACCCAAACATCAA 661
QY 546 TGACGACATGCTGTGTGCGG-----GAACACCCGGAAGGACTATGCCAGGGGACTTC 599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 662 GATGACATGCTGTGTGCGGCTTCGAGAGGAGGCAAGAGATGCTCGAAGGGGCACTC 721
QY 600 CGAGGGGCCCCGTGTGTGCAAGTGAATGCACTGGCTGCAAGGGGGGTGTCAGCTG 659
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 722 GGGCGGCCCCCTGTGTGTGCTGTGAGTGTGCTGTGAGGGGGGGGTATATGCTG 781
QY 660 GGGCGAGGGCTGTGCGCACGCCCAACCGGCTGGCATCTACACCGGTGTGCTACTACTT 719
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 782 GGGTGTAGGGCTGTGCGCCGACAGACCGCCAGGTGTCTACATCCGTGTGACCGCCACCA 841
QY 720 GGACTGTGATCCACACTATGTGTCCTCCCAAAAGCCGT 754
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 842 CAAGTGTGATCATCGATCATCCCAAACTGCAGT 876
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
US-10-028-072-221
; Sequence 221, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang

```

APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gunney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang
TITLE OR INVENTION:
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028, 072
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836
PRIOR FILING DATE: 1997-09-24
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062814
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063561
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063704
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063735
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03

PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072320
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579


```

; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086414
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086430
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088730
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088741
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 19/98-06-11
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090538
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07

```

Query Match 26.1%; Score 201.4; DB 9; Length 1129;
Best Local Similarity 57.4%; Pred. No. 1.2e-38;
Matches 433; Conservative 0; Mismatches 301; Indels 21; Gaps 3;

```

QY 12 GAAAGATCGTGGGGGTGAGAGGCCCCCAAGAGCAGTGGCCCTGCGAGTGGCT 71
DB 131 GAACCAATGTTGGGGGAGGACGACGAGAGGGGCGAGTGGCCCTGCGAGTGGCT 190
QY 72 GAGAGTCAGGCGCCATGATGATCTTCTGCGGGGCTCCCTCATCCACCAGT 131
DB 191 -----CGAGCGCAAGGAAAGCCATTCTGCGGGGCGACGCTATGCGGAGCA 241
QY 132 GGTGCTGACCCCGCGGCTGCGTGGGACCGGACGTCAAGATCTGGCC-----GCCCT 185
DB 242 GGTCTGAGGCGTGGCTGCTCCGCAACACTCTGAGAGCTCCCTGATACAGGTCT 301
QY 186 CAGGGGCAATGCGGGGAGCAGCACTCTACTACAGGACCAAGCTGCTCCGCTGAG 245
DB 302 GCTGGGGGCAAGGACACTGTGTAGGCGGGACACAGCTATGTATGCCGGGTGAGCA 361
QY 246 GATCATCTGACCCACAGATCTTACACCGCCAGATCGAGGAGGACATGCGCTGTGA 305
DB 362 GGTGAGAGCAACCCCTTACCAAGGGGCAAGGCGCTCCAGCGCTGAGTGGCTGTGA 421
QY 306 GCTGAGAGGAGCGGTGAGGTCTCCAGCCAGTCCACAGGCTACCTGCCCCCTGCTTC 365
DB 422 GCTGAGGACCAAGTGCCTTCAACCAATTCATCTCCCGGTGCTGCTGACCCCTTC 481

```

```

QY 366 AGAGACTTCCCCCGGGAGTCCGTGCGGTCACATGGCTGGGCGATGTGACATGA 425
DB 482 GGTGATCTTTGAGAGGGGATTAAGTCTGCTGCTGCTGCTGGGACACCCAGTGA 541
QY 426 TGAGGCGCTCCGACCGCATTTCTCTGAGAGGAGTGAAGTCCCATATGGAACCA 485
DB 542 AGACTCTCGCCGGAACCGCGGATCTGAGAACTCGCTGTGCCATCATGACACACC 601
QY 486 CATTTGTAGCGAAATATACCACTTGGCGCTACACGGGAGACGACGTCCGATGTC 545
DB 602 CAAGTCAACCTGCTCTACAGCAAGACCGAGTTTGGCTACCAACCCAAAACATCA 661
QY 546 TCAGCACATGCTGTGTCGCGG-----GAACACCGGAGAGGACTATGACGAGGACT 599
DB 662 GAATGACATGCTGTGCGCGCGCTTGAGAGAGGCAAGAAAGATGCTCTGCAAGG 721
QY 600 CGAGGGGCGCTGTGTGCAAGTGAATGACACTGCTGCAAGGCGGCGTGTGACGTG 659
DB 722 GGGCGGCGCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 781
QY 660 GGGCGAGGCTGTGCGCGGACCGCAACCGGCTGTCATCTACACCGCTGCTACTACT 719
DB 782 GGTGAGGCGTGTGCGCGGACCGCAACCGGCTGTCATCTGCTGCTGCTGCTGCT 841
QY 720 GGACTGATCCACCACTATGTCCCAAAAAGCGGT 754
DB 842 CAACTGATCCATCGGATCATCCCAACTGCAGT 876

```

```

RESULT 6
US-10-041-006A-8
; Sequence 8, Application US/10041006A
; Patent No. US20020168754A1
; GENERAL INFORMATION:
; APPLICANT: Andrade-Gordon, Patricia
; APPLICANT: Darrow, Andrew
; APPLICANT: Qi, Jian-shen
; TITLE OF INVENTION: DNA encoding the novel human serine
; FILE REFERENCE: ORT-1032
; CURRENT APPLICATION NUMBER: US/10/041.006A
; CURRENT FILING DATE: 2002-01-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1130
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene of
; OTHER INFORMATION: Protease T in a zymogen activation vector
US-10-041-006A-8

```

Query Match 25.1%; Score 193.8; DB 9; Length 1130;
Best Local Similarity 56.7%; Pred. No. 7.5e-37;
Matches 429; Conservative 0; Mismatches 307; Indels 21; Gaps 3;

```

QY 10 GAGAAAAGATCGTGGGGGTGAGAGGCCCCAGAGCAAGTGGCCCTGCGAGTGA 69
DB 157 GATGACAAAGATGTTGGGGGTATGCTTAGAGAGGGGAGTGGCCCTGCAAGTCA 216
QY 70 CTGAGAGTCCAGGCGCCATACTGATGATCTTCTGCGGGGCTCCCTCATCCACCC 129
DB 217 AT-----CGAGGCAAGCGGAAGCCACTTCTGCGGGGAGGCTCTATCCGAG 267
QY 130 TGGGTGCTGACCGCGCGCGCTGCGTGGAGGAGCGAGTCAAGATCTGA-----CCGCC 183
DB 268 TGGGTCTTACCGGCTGCGGCTGCTTCCGACACACTCTGAACAGTCCCTGTACAG 327
QY 184 CTGAGGTCCAACTGCGGGAGGAGCACTCTACTACAGGACAGTGTGCGCGGTGAG 243
DB 328 CTGCTGGGGGCAAGGAGCTAGTGCAGCGGGAGCACACAGCTATGTATGCGCGGTGAG 387

```


Matches 370; Conservative 0; Mismatches 234; Indels 18; Gaps 2;

```
OY 141 CGCCGCGGCTGCGGAGCGGACGTCAGAGATCTGGCCGCTCAGGGTGCACCTCG 200
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 612 CCCCCCGCCCTGCAGGTCAGGAGAGATTGGAGCTTCCCGTTTAAAGTGCAGCTGG 671
OY 201 GGAGCAGACCTCTACTACAGAGACGCTCTGCGGTCAGCAGGATCATCTGACCC 260
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 672 GCAGCTGAGCTCTATGAGAGACGACGAGAGAGTGGTGTGATCTGTCCTGACCC 731
OY 261 ACAGTTCTACA-----CCGCCAGATGGAGCCGACATCCGCTGCTGAGACTGA 311
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 732 CCAGTACAGACAGAGACCTGTCTGCCAGGCGGTGCAGCATCGCCGCTGAACCTGA 791
OY 312 GGAGCGGCTGAGGCTCTGACGACGCTGACAGGCTGACCCGCTGCTGAGAGAC 371
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 792 GGGCCGCTGCGGCTGTGAGCTATCCACCGGCTCTCGCTCCGCTGCTCTCTGA 851
OY 372 CTTCGCCCGGGGATCGGCTGCTGAGCTGAGGCGCATGTGACAAATGATGAGCG 431
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 852 CGTGCCCTCGGGGAAGACCTGCTGAGGAGCGGCTGAGGCTGATTTGAGCGAGAACT 911
OY 432 CCTCCACCGGCTATCTCTGAGACAGTGAAGTCCCATATGAGAAACCACTTTG 491
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 912 ACTGCTCTGCGCCCTCAGCTTGTGGAGGCGAGCTGAAGTCAAGAGCAAGTCTCTG 971
OY 492 TGACGCAAAATACAC-----CTTGGCGGCTACAGGGAGAGCAAGTCCGCTG 542
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 972 TACACGACCTGTGCGCCGCTTCTTCCACACACTGAGCGGTTTGAAGCGCTAT 1031
OY 543 CCGTGACGACATGTGTGTGCGGAGACACCGGAGGACTCATGCCAGGAGCTCGG 602
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1032 CAAGAGCAGATGTGTGTGCGGAGGAGGAGAACACAGGCTCTGCGGAGGAGACAG 1091
OY 603 AGGCGCTGCTGTGTCAGAGTGAATGACACTGCTGACAGCGGCGTGTGACCTGG 662
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1092 GGGCCCTCTGTGTCAGGCGGAATGACCTGGGTCAGGAGGAGTGTGAGCTGGGG 1151
OY 663 CGAGGCTGTGCTGCCAGCCGACCGGCTGAGCATACCGCTGACCTACTACTGA 722
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1152 CAACTCTGCGGCTTGTGCGGCTATCCCGGCTATGACCCGCTGACGAGCTACGTC 1211
OY 723 CTGATCCACGACTATGTCCT 744
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1212 CTGATCCGCGAGCAGCTGCCCC 1233
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 9

```
US-10-041-400A-1
: Sequence 1, Application US/10041400A
: Patent No. US20020110895A1
: GENERAL INFORMATION:
: APPLICANT: Darrow, Andrew
: APPLICANT: Andrade-Gordon, Patricia
: APPLICANT: Qi, Jenson
: TITLE OF INVENTION: DNA Encoding the Human Serine
: FILE REFERENCE: ORT-1031
: CURRENT APPLICATION NUMBER: US/10/041,400A
: CURRENT FILING DATE: 2002-01-08
: PRIOR APPLICATION NUMBER: US/09/387,375
: PRIOR FILING DATE: 1999-08-31
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1613
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-041-400A-1
```

Query Match 24.4%; Score 187.8; DB 12; Length 1613;
Best Local Similarity 56.9%; Pred. No. 2,le-35;
Matches 414; Conservative 0; Mismatches 292; Indels 21; Gaps 3;

```
OY 17 GAATGTCGGGGTTCAGAGAGCCGCCAGAGCAAGTGGCCCTGGCAGGTGAGAG 76
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 175 GGATGTTGGGGGCCGGATGATGCGGAGAGAGATGGCCCTGGAGGCGAGATCCAGC 234
OY 77 TCCAGGCCCATATCTGATGATGCACTTCTGCGGGGCTCTCCATATCCACCCAGTGGTGC 136
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 235 ATCTCGGGGCA-----CACGTGTGCGGGGGTCTCATATCCGCCAGTGGGTGC 285
OY 137 TGACCGCGCGGCTGCTGCGGAGCCGAGACGTCAAG-----GATCGGCCGCCCTCAGGG 190
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 286 TGACAGCGGCGCACTGCTTCCCAAGAGGAGCACTCCAGCTGAGTACCGCGTGCCTGG 345
OY 191 TGCAACTGCGGGAGCAGACCTCTACTACAGAGACCACTCTGCGGCTGACAGATCA 250
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 346 GGGCGCTGCTGTGGCTTCACACTGCGCCCGACGCTCTCGGTGCGGCTGGAGCGGTGC 405
OY 251 TGTGACCCACAGTTCTACACCGGCCAGATCGAGCGGAGATCGCCCTGCTGAGCTGG 310
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 406 TGTGCCCCCGACTACTCCAGAGACGCGGCGCGGCGGAGCTGGCACTGCTGACGTGC 465
OY 311 AGGACCGGCTGAAGTCTTCACGCCACGTCCACAGCGTCACTGACCCCTGCTGACAGA 370
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 466 GTGCGCGGCTGCGCTGAGCGCTGCGCTCAACCGCTGCTGCGCCGCTGCGCGCGCC 525
OY 371 CTTTCCCGCGGGGATGCGCTGCTGCTGCTGCTGCTGCGGCGATGTGACATGATGAGC 430
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 526 GCGCGCGCGCGGCGACACCATGCGGCTCACCGGCTGCGGAGACCTCCGCCAGAGAGTGC 585
OY 431 GCTTCCACCGCCATTTCTCTGTAAGCAGGTGAAGTGCATTAATGAGAAACCACTTT 490
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 586 CCGTCCAGAGATGGCGACTCGGTACAAAGAGATAAGGTCGCGCTGTGATCGGCACTCT 645
OY 491 GTGACGCAAAATACCACTTGGCGCTTACAGGAGAGAGAGCTCGCATCTGCTGACG 550
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 646 GCGAGGCGCTTACCACTGTGGCGGAGCGTGCACGAGCTGAGAGCAATTGTGCTGCTG 705
OY 551 ACATGCTGTGCGCGGGAACACCGG-----AGGACTCATGTCAGAGGCGCATCGGAG 604
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 706 GAGCTCTGTGTCGCGGCTACCCCGCAGGGCGACAGAGACCTGCGAGGCTATTTCTGGGG 765
OY 605 GGGCCCTGCTGTCAAGGTGAATGAGCAGCTGCTGACGCGGCGGTGTGCTAGCTGGGCG 664
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 766 GACCTCTGACCTGCTGCAAGTCTGGGAGCTGGGTCTGTGTTGGGTGTGAGTGGGCA 825
OY 665 AGGCTGTGCGGACGCGCAAGCGGCTGGCATCTACACCGCTGCTCACTACTTGGACT 724
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 826 AGGTTGTGCTTGCCTGCCCAACGCTCCAGGGGTCTACACCAAGTGTGCGCACATATAGCCCT 885
OY 725 GGATCCA 731
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 886 GGATTCA 892
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 10

```
US-10-041-264A-1
: Sequence 1, Application US/10041264A
: Patent No. US2002014246A1
: GENERAL INFORMATION:
: APPLICANT: Darrow, Andrew
: APPLICANT: Andrade-Gordon, Patricia
: APPLICANT: Qi, Jenson
: TITLE OF INVENTION: DNA Encoding the Human Serine
: FILE REFERENCE: ORT-1031
: CURRENT APPLICATION NUMBER: US/10/041,264A
: CURRENT FILING DATE: 2002-01-08
: PRIOR APPLICATION NUMBER: US/09/387,375
: PRIOR FILING DATE: 1999-08-31
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1613
```

TYPE: DNA
ORGANISM: Homo sapiens
US-10-041-264A-1

Query Match 24.4%; Score 187.8; DB 12; Length 1613;
Best Local Similarity 56.9%; Pred. No. 2.1e-35;
Matches 414; Conservative 0; Mismatches 292; Indels 21; Gaps 3;

QY 17 GAATGCTGGGGGCTACAGAGGCCCCCAGACCAAGTGGCCCTGGAGGTGAGCCGGAAG 76
DB 175 GGATGCTGGGGGCGCGGATGCGCGGAGCAGAGTGGCCCGTGAGGCGAGATCCAGC 234
QY 77 TCCAGGCGCCATGATGATGATCTTCTGCGGGGGCTCCCTATCCACCCCGAGTGGTGC 136
DB 235 ATCTGGGGGCA-----CAGTGTGCGGGGGGTGCTATGCCCCCAGTGGGTGC 285
QY 137 TGACCGCGCGGCGTGGCTGGAGCCGACGTCAG-----GATCTGGCCGCTCAGGG 190
DB 286 TGACAGCGCGGCGACTGCTTCCCAAGAGGCACTGCCAGTGAATCCGCGTGCCTGG 345
QY 191 TGCAACTGCGGGAGCAGCCTCTACTACAGAGCACTGCTGCGCGGTCAAGCAGATCA 250
DB 346 GGGCGCTGCTGTGGGCTCCACCTTGCGCCGACGCTTCGCTGCTGCGAGCGGTGC 405
QY 251 TGTGACACCAAGTTCTACACCGCCAGATCGAGCGAGATCGCCCTGCTGAGCTGG 310
DB 406 TGTGCGCCCGGAGTACTCCGAGAGCGGGGCGCGGCGGCTGCTGCTGCTGCTGCTG 465
QY 311 AGGAGCGCGGTGAAGTCTTCACGCCACGTCACACGCTACCTGCCCCCTGCTCAGAG 370
DB 466 GTGCGCGCGTGCCTGAGCGCTGCGTCAACCGCTGCTGCTGCTGCTGCTGCTGCTG 525
QY 371 CTTTCCCCCGGGGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 430
DB 526 GCGCGCGCGCGGACACACATGCGCGGTGCTGCGGTGCGGTGCGGTGCGGTGCGGTG 585
QY 431 GCTTCCACCGCCATTTCTCTGGAAGAGGTGAAGTCCCATTAATGAAAAACCATTT 490
DB 586 CCTTCCAGATGAGGCGCGCTACAAAGAGTGAAGGTGCGGTGCTGCTGCTGCTGCTGCT 645
QY 491 GTGACGCAAAATACCACTTGGCGCTACACGGAGAGAGAGCTCGCATGCTCGCTGAC 550
DB 646 GCGAGCGCGCTTACACAGTGGGCGGAGCTGCCCCAGGCTGAGGCAATTTGCTGCTG 705
QY 551 ACATGCTGTGCGCGGACACCGG-----AGGACTCATGCGAGGCGCATCGGAG 604
DB 706 GGAAGTCTGTGCGCGGCTACCCCGGAGGCGACAGAGAGCGCTGCGAGGCTATTTGG 765
QY 605 GCGCCCTGCTGCAAGTGAATGAGCAGCTGCTGCGAGCGGGCGTGTGACAGTGGGCG 664
DB 766 GACCTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 825
QY 665 AGGCTGTGCGCGGACCGGCGCTGCGATCTACACCGGTGCTACCTACTTGGACT 724
DB 826 AGGCTGTGCGCGGCTGCGCAACCGTCCAGGGGTCTACACCAAGTGTGCGCATATAG 885
QY 725 GGATCCA 731
DB 886 GGATTCA 892

RESULT 11
US-10-042-091A-1
Sequence 1, Application US/10042091A
Patent No. US20020142447A1
GENERAL INFORMATION:
APPLICANT: Darrow, Andrew
APPLICANT: Andrade-Gordon, Patricia
APPLICANT: Qi, Jensen
TITLE OF INVENTION: DNA Encoding the Human Serine
FILE REFERENCE: ORT-1031
CURRENT APPLICATION NUMBER: US/10/042.091A

CURRENT FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US/09/387,375
PRIOR FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1
LENGTH: 1613
TYPE: DNA
ORGANISM: Homo sapiens
US-10-042-091A-1

Query Match 24.4%; Score 187.8; DB 12; Length 1613;
Best Local Similarity 56.9%; Pred. No. 2.1e-35;
Matches 414; Conservative 0; Mismatches 292; Indels 21; Gaps 3;

QY 17 GAATGCTGGGGGCTACAGAGGCCCCCAGACCAAGTGGCCCTGGAGGTGAGCTGAGAG 76
DB 175 GGATGCTGGGGGCGCGGATGCGCGGAGCAGAGTGGCCCGTGAGGCGAGATCCAGC 234
QY 77 TCCAGGCGCCATGATGATGATCTTCTGCGGGGGCTCCCTATCCACCCCGAGTGGTGC 136
DB 235 ATCTGGGGGCA-----CAGTGTGCGGGGGGTGCTATGCCCCCAGTGGGTGC 285
QY 137 TGACCGCGCGGCGTGGCTGGAGCCGACGTCAG-----GATCTGGCCGCTCAGGG 190
DB 286 TGACAGCGCGGCGACTGCTTCCCAAGAGGCACTGCCAGTGAATCCGCGTGCCTGG 345
QY 191 TGCAACTGCGGGAGCAGCCTCTACTACAGAGCACTGCTGCGCGGTCAAGATCA 250
DB 346 GGGCGCTGCTGTGGGCTCCACCTTGCGCCGACGCTTCGCTGCTGCTGCTGCTGCTGCTG 405
QY 251 TGTGACACCAAGTTCTACACCGCCAGATCGAGCGAGATCGCCCTGCTGAGCTGG 310
DB 406 TGTGCGCCCGGAGTACTCCGAGAGCGGGGCGCGGCGGACTGCTGCTGCTGCTGCTG 465
QY 311 AGGAGCGCGGTGAAGTCTTCACGCCACGTCACACGCTACCTGCCCCCTGCTCAGAG 370
DB 466 GTGCGCGCGTGCCTGAGCGCTGCGTCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 525
QY 371 CTTTCCCCCGGGGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 430
DB 526 GCGCGCGCGCGGACACACATGCGCGGTGCTGCGGTGCGGTGCGGTGCGGTGCGGTG 585
QY 431 GCTTCCACCGCCATTTCTCTGGAAGAGGTGAAGTCCCATTAATGAAAAACCATTT 490
DB 586 CCTTCCAGATGAGGCGCGCTACAAAGAGTGAAGGTGCGGTGCTGCTGCTGCTGCTGCT 645
QY 491 GTGACGCAAAATACCACTTGGCGCTACACGGAGAGAGAGCTCGCATGCTCGCTGAC 550
DB 646 GCGAGCGCGCTTACACAGTGGGCGGAGCTGCCCCAGGCTGAGCGCAATTTGCTGCTG 705
QY 551 ACATGCTGTGCGCGGACACCGG-----AGGACTCATGCGAGGCGCATCGGAG 604
DB 706 GGAAGTCTGTGCGCGGCTACCCCGGAGGCGACAGAGAGCGCTGCGAGGCTATTTGG 765
QY 605 GCGCCCTGCTGCAAGTGAATGAGCAGCTGCTGCGAGCGGGCGTGTGACAGTGGGCG 664
DB 766 GACCTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 825
QY 665 AGGCTGTGCGCGGACCGGCGCTGCGATCTACACCGGTGCTACCTACTTGGACT 724
DB 826 AGGCTGTGCGCGGCTGCGCAACCGTCCAGGGGTCTACACCAAGTGTGCGCATATAG 885
QY 725 GGATCCA 731
DB 886 GGATTCA 892

RESULT 12
US-09-900-754-1
Sequence 1, Application US/09900754
Patent No. US2002002654A1
GENERAL INFORMATION:

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2003, 03:51:33 : Search time 56 Seconds
(without alignments)
4222.282 Million cell updates/sec

Title: US-09-598-982-20

Perfect score: 771

Sequence: 1 ggagccctcgagaaagaat.....cgtgaagcgccgcctcgt 771

Scoring table: IDENTITY_NUC

Gapop 10.0', Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_NA.*
2: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PC70S.COMB.seq:*
7: /cgn2_6/ptodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	764.6	99.2	771	4	US-09-079-970A-4 Sequence 4, Appl
2	731	94.8	1128	2	US-09-016-366A-20 Sequence 20, Appl
3	731	94.8	1128	2	US-08-978-404B-15 Sequence 15, Appl
4	731	94.8	1137	2	US-09-016-366A-18 Sequence 18, Appl
5	731	94.8	1137	2	US-08-978-404B-13 Sequence 13, Appl
6	728.6	94.5	735	4	US-09-079-970A-1 Sequence 1, Appl
7	728.2	94.2	1081	2	US-09-016-366A-22 Sequence 22, Appl
8	726.2	94.2	1081	2	US-08-978-404B-17 Sequence 17, Appl
9	686.2	89.0	1154	2	US-09-016-366A-16 Sequence 16, Appl
10	686.2	89.0	1154	2	US-08-978-404B-11 Sequence 11, Appl
11	477.2	61.9	1219	2	US-08-978-404B-7 Sequence 7, Appl
12	469.6	60.9	1108	2	US-09-016-366A-14 Sequence 14, Appl
13	469.6	60.9	1108	2	US-08-978-404B-20 Sequence 20, Appl
14	463	60.1	1031	2	US-08-978-404B-1 Sequence 1, Appl
15	445.6	57.8	1103	2	US-09-016-366A-24 Sequence 24, Appl
16	444	57.6	1097	2	US-08-978-404B-4 Sequence 4, Appl
17	371	48.1	2259	2	US-08-845-998-3 Sequence 3, Appl
18	371	48.1	2259	3	US-09-206-537-3 Sequence 3, Appl
19	371	48.1	2259	4	US-09-430-854-3 Sequence 3, Appl
20	369.4	47.9	2218	4	US-08-845-998-5 Sequence 5, Appl
21	369.4	47.9	2218	4	US-09-206-537-5 Sequence 5, Appl
22	369.4	47.9	2218	4	US-09-430-854-5 Sequence 5, Appl
23	271.8	35.3	1095	2	US-08-978-404B-9 Sequence 9, Appl
24	201.4	26.1	1110	4	US-09-386-653A-1 Sequence 1, Appl
25	193.8	25.1	1130	4	US-09-386-653A-8 Sequence 8, Appl
26	153	19.8	3757	2	US-09-016-366A-13 Sequence 13, Appl
27	153	19.8	3757	2	US-08-978-404B-19 Sequence 19, Appl

28	143.8	18.7	2397	2	US-08-978-404B-2 Sequence 2, Appl
29	140.6	18.2	1142	4	US-09-386-642-8 Sequence 8, Appl
30	140.6	18.2	1169	4	US-09-386-642-7 Sequence 7, Appl
31	131.2	17.0	1430	4	US-09-386-629-1 Sequence 1, Appl
32	129.2	16.8	1166	4	US-09-386-629-2 Sequence 2, Appl
33	115.4	15.0	1081	4	US-09-008-271A-15 Sequence 15, Appl
34	115.4	15.0	1103	4	US-09-386-642-59 Sequence 59, Appl
35	112.8	14.6	2413	4	US-09-518-046-1 Sequence 1, Appl
36	108.6	14.1	1109	4	US-09-088-651-6 Sequence 6, Appl
37	107.6	14.0	1109	4	US-09-088-651-6 Sequence 1, Appl
38	103.6	13.4	1386	2	US-08-756-506-3 Sequence 3, Appl
39	103.6	13.4	1755	6	5225537-1 Patent No. 5225537
40	102.8	13.3	2416	4	US-09-261-416-1 Sequence 1, Appl
41	100.8	13.1	654	4	US-09-280-116-57 Sequence 57, Appl
42	98.4	12.8	2544	4	US-09-518-046-3 Sequence 3, Appl
43	95.6	12.4	1155	4	US-09-163-951-15 Sequence 15, Appl
44	95.6	12.4	1240	4	US-09-163-951-14 Sequence 14, Appl
45	94.2	12.2	1387	6	5270178-1 Patent No. 5270178

ALIGNMENTS

```
RESULT 1
US-09-079-970A-4
; Sequence 4, Application US/09099970A
; Patent No. 6274366
; GENERAL INFORMATION:
; APPLICANT: Mafilt, Mark A.
; APPLICANT: Niles, Andrew L.
; APPLICANT: Haak-Frendscho, Mary
; TITLE OF INVENTION: Enzymatically-Active Recombinant Human
; TITLE OF INVENTION: Beta-Tryptase and Method of Making Same
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Intellectual Property Department
; STREET: 8000 Excelsior Drive, Suite 401
; CITY: Madison
; STATE: WISCONSIN
; COUNTRY: U.S.A.
; ZIP: 53717-1914
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079, 970A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Leone, Joseph T.
; REGISTRATION NUMBER: 37,170
; REFERENCE/DOCKET NUMBER: 34506.073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 831-2100
; TELEFAX: (608) 831-2106
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7..753
; FEATURE:
; NAME/KEY: misc_signal
```


QY	557	TGTTGGCGGGAAACACCCGGAGGAGCACTCAATGCCAGAGGGCGACCTCGAGAGGGCCCCCTGGTGT	616
Db	627	TGTGTGCGGGGAATACCCCGGAGGGAGCACTCAATGCCAGAGGGCGACCTCGAGAGGGCCCCCTGGTGT	686
QY	617	GCAAGTCGTAATGAGCACCCTGGCTGGAGGCGGGGCGTGTGCACACTGGGGGCGAGAGCGCTGTGTGCC	676
Db	687	GCAAGTGTGAATGAGCACCCTGGCTGGAGGCGGGGCGTGTGTGCACACTGGGGGCGAGAGCGCTGTGTGCC	746
QY	677	AGCCCAACCGGGCTGGCATCTACACCCCGTGCACCTACTACTTGGATGGATGCACACACT	736
Db	747	AGCCCAACCGGGCTGGCATCTACACCCCGTGCACCTACTACTTGGATGGATGCACACACT	806
QY	737	ATGTCCCCAAAAGCCGTGAAGCGGCC	763
Db	807	ATGTCCCCAAAAGCCGTGAAGCGGCC	833

RESULT 3
US-08-978-404B-15
; Sequence 15, Application US/08978404B

GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.
TITLE OF INVENTION: MAST CELL PROTEINASE THAT CLEAVES
TITLE OF INVENTION: FIBRINOGEN
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,404B
FILING DATE: 25-NOV-97
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/032,354
FILING DATE: 04-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7090
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1128 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-978-404B-15

Query Match	94.8%;	Score 731;	DB 2;	Length 1128;
Best Local Similarity	98.7%;	Pred. No. 6e-157;		
Matches 737;	Conservative	0;	Mismatches 10;	Indels 0;
			Gaps	0

Qy	17	GAATGTCGTGGGGGGTCACAGAGGCCCCCAGAGGACCAAGTGGCCCTGGCAGAGTCAAGCTTGAGAG	76
Db	87	GCATGTTGGGGGTCAGAGAGCCCCCAGAGCAGGTGGCCCTGGCAGAGTCAAGCTTGAGAG	146
Qy	77	TCCAGGGCCCACTACTGATGATGACTTTCGCGGGGGCTCCCTCATCATCAACCCCAAGTGGGTGC	136
Db	147	TCCAGGGCCCAATACGTGATGATGACTTTCGCGGGGGCTCCCTCATCATCAACCCCAAGTGGGTGC	206

Oy	137	TGACCGCGGCGGCTGCTGGGAGACGGGACGTCAGATCTCGCCGCCCTCAGGGTGCAC	196
Oy	137		
Db	207	TGACCGCAGCGCCACTGCTGGGAGCGGACGTCGAAGATCTGGCCGCTTAGGGTGCAC	266
Oy	197	TGCGGGAGACACACTCTCTACTACCAGGACGAGCTGCTGCCGGTCAAGAGATCATCTGC	256
Db	267	TGCGGGAGACACACTCTCTACTACAGGACCAAGCTGCTGCCGGTCAAGAGATCATCTGC	326
Oy	257	ACCCACAGTTCTACACCGGCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGAGC	316
Db	327	ACCCACAGTTCTACACCGGCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGAGC	386
Oy	317	CGGGAAGGTCCTCAGCAGCTCCACACGGGTCACCTCGCCCGCTCGCTCAGACCTTTC	376
Db	387	CGGGAAGGTCCTCAGCAGCTCCACACGGTCACTCCCTGCTCGCTCAGACCTTTC	446
Oy	377	CGCGGGGATGCGTGTGCTGCTACTGCTGGGGCGATGTGGACAATGATAGCGCTTC	436
Db	447	CGCGGGGATGCGTGTGCTGCTACTGCTGGGGCGATGTGGACAATGATAGCGCTTC	506
Oy	437	CACGGCCATTTCTCTGAGAGAGGTGAAGTCCCATTAATGGAAACACATTTGTACG	486
Db	507	CACGGCCATTTCTCTGAGAGAGGTGAAGTCCCATTAATGGAAACACATTTGTACG	566
Oy	497	CAAAATACACACTTGGGCGCTACACGGGAGACGACGTCGCCATGCTCGTAGACACATGC	556
Db	567	CAAAATACACACTTGGGCGCTACACGGGAGACGACGTCGCCATGCTCGTAGACACATGC	626
Oy	557	TGTGTGCCGGGAACACCCGGAGGACTTCATCCAGGGCGACTCCGGAGGGCCCTGCTGT	616
Db	627	TGTGTGCCGGGAACACCCGGAGGACTTCATCCAGGGCGACTCCGGAGGGCCCTGCTGT	686
Oy	617	GCAAGTCAATGGGACCGCTGGTAGGGGGCGGAGTCAGCTGGGGGAGAGGCTGTCC	676
Db	687	GCAAGTCAATGGGACCGCTGGTAGGGGGCGGAGTCAGCTGGGGGAGAGGCTGTCC	746
Oy	677	AGCCCAACCGGCTGGCATCTACACCGGTCTCACCTACTACTTGAATGATCCACACT	726
Db	747	AGCCCAACCGGCTGGCATCTACACCGGTCTCACCTACTACTTGAATGATCCACACT	806
Oy	737	ATGTCTCCCAAAAAGCGGTGAAGCGGCC	763
Db	807	ATGTCTCCCAAAAAGCGGTGAAGCGGCC	833

RESULT 4
US-09-016-366A-18

1 GENERAL INFORMATION:
2 APPLICANT: Stevens, Richard L.
3 APPLICANT: Huang, Chitu
4 TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
5 TITLE OF INVENTION: INHIBITORS
6 NUMBER OF SEQUENCES: 65
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
9 STREET: 600 Atlantic Avenue
10 CITY: Boston
11 STATE: MA
12 COUNTRY: U.S.A.
13 ZIP: 02210-2211
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Diskette
16 COMPUTER: IBM Compatible
17 OPERATING SYSTEM: DOS
18 SOFTWARE: FASTSEQ for Windows Version 2.0
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/09/016,366A
21 FILING DATE: January 30, 1998
22 CLASSIFICATION: 530
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: 60/037,090

QY 77 TCACGGCCACTAGTATGATGCTTCTGCGGGGCTCCCTCATCCACCCAGTGGGTGC 136
Db 126 TCACGGCCACTAGTATGATGCTTCTGCGGGGCTCCCTCATCCACCCAGTGGGTGC 185
QY 137 TGACCGCCGGCGGCTGCTGGGACCGGACGTCAAGATCTGGCGCCCTCAGGTGCAAC 196
Db 186 TGACCGCCGGCGGCTGCTGGGACCGGACGTCAAGATCTGGCGCCCTCAGGTGCAAC 245
QY 197 TGCGGGACGACACCTCTACTACGAGACAGCTGCTGCGGTGACGAGATCATGCTGC 256
Db 246 TGCGGGACGACACCTCTACTACGAGACAGCTGCTGCGGTGACGAGATCATGCTGC 305
QY 257 ACCACAGCTTCTACACCGCCAGATCGGAGCGGACATCGCTGCTGAGCTGAGAGAGC 316
Db 306 ACCACAGCTTCTACACCGCCAGATCGGAGCGGACATCGCTGCTGAGCTGAGAGAGC 365
QY 317 CGGTGAAGTCTCCAGGACGCTCCACAGGCTACCCCTGCTGCTGCTGCTGCTGCTGCT 376
Db 366 CGGTGAAGTCTCCAGGACGCTCCACAGGCTACCCCTGCTGCTGCTGCTGCTGCTGCT 425
QY 377 CCGCGGGGATCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436
Db 426 CCGCGGGGATCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 485
QY 437 CACCGCCATTTCTCTGAGAGAGGTGAAGTCCCATATGAGAAACCAATTTGTGAGC 496
Db 486 CACCGCCATTTCTCTGAGAGAGGTGAAGTCCCATATGAGAAACCAATTTGTGAGC 545
QY 497 CAAATATCCACTTGGGCTTACACGAGGAGACGATCCGATCGCTGCTGAGCATGTC 556
Db 546 CAAATATCCACTTGGGCTTACACGAGGAGACGATCCGATCGCTGCTGAGCATGTC 605
QY 557 TGTGTGGCGGGAACCGCGGAGGAGCTATGCGAGGCGACGCGGCGGCGGCGGCGT 616
Db 606 TGTGTGGCGGGAACCGCGGAGGAGCTATGCGAGGCGACGCGGCGGCGGCGGCGT 665
QY 617 GCAAGGTGAATGGACACTGCTGACGAGGCGGCTGTGCTGCTGCTGCTGCTGCTGCT 676
Db 666 GCAAGGTGAATGGACACTGCTGACGAGGCGGCTGTGCTGCTGCTGCTGCTGCTGCT 725
QY 677 AGCCCAACCGGCTTGGCATCTACACCCGTGCTACTACTTGGAGCTGATCCACACT 736
Db 726 AGCCCAACCGGCTTGGCATCTACACCCGTGCTACTACTTGGAGCTGATCCACACT 785
QY 737 ATGTCCCAAAAAGCGCTGAAGGCGC 763
Db 786 ATGTCCCAAAAAGCGCTGAAGGCGC 812

RESULT 9
US-09-016-366A-16
Sequence 16, Application US/09016366A
Patent No. 5955431
GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.
APPLICANT: Huang, Chifu
TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
TITLE OF INVENTION: INHIBITORS
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,366A

FILING DATE: January 30, 1998
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/037,090
FILING DATE: 05-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7093
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1154 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-016-366A-16

Query Match 89.0%, Score 686.2; DB 2; Length 1154;
Best Local Similarity 94.3%; Pred. No. 8,2e-147;
Matches 712; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 17 GAATGTCGGGGGTTCAGGAGGCCCCCAGAGCAAGTGGCCCTGGAGGTGAGCCTGAGAG 76
Db 106 GATGTCGGGGGTTCAGGAGGCCCCCAGAGCAAGTGGCCCTGGAGGTGAGCCTGAGAG 165
QY 77 TCACGGCCACTAGTATGATGCTTCTGCGGGGCTCCCTCATCCACCCAGTGGGTGC 136
Db 166 TCACGGCCACTAGTATGATGCTTCTGCGGGGCTCCCTCATCCACCCAGTGGGTGC 225
QY 137 TGACCGCCGGCGGCTGCTGGGAGCCGAGCTCAAGATCTGGCGCCCTCAGGTGCAAC 196
Db 226 TGACCGCCGGCGGCTGCTGGGAGCCGAGCTCAAGATCTGGCGCCCTCAGGTGCAAC 285
QY 197 TGCGGGACGACACCTCTACTACGAGACAGCTGCTGCGGTGACGAGATCATGCTGC 256
Db 286 TGCGGGACGACACCTCTACTACGAGACAGCTGCTGCGGTGACGAGATCATGCTGC 345
QY 257 ACCACAGCTTCTACACCGCCAGATCGGAGCGGACATCGCCCTGCTGAGCTGAGAGAGC 316
Db 346 ACCACAGCTTCTACATCATTCAGACTGGAGGAGATGCGCTGCTGAGCTGAGAGAGC 405
QY 317 CGGTGAAGTCTCCAGGACGCTCCACAGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCT 376
Db 406 CGGTGAAGTCTCCAGGACGCTCCACAGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCT 465
QY 377 CCGCGGGGATCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436
Db 466 CCGCGGGGATCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 525
QY 437 CACCGCCATTTCTCTGAGAGAGGTGAAGTCCCATATGAGAAACCAATTTGTGAGC 496
Db 526 CACCGCCATTTCTCTGAGAGAGGTGAAGTCCCATATATGAGAAACCAATTTGTGAGC 585
QY 497 CAAATATCCACTTGGGCTTACACGAGGAGACGATCCGATCGCTGCTGAGCATGTC 556
Db 586 CAAATATCCACTTGGGCTTACACGAGGAGACGATCCGATCGCTGCTGAGCATGTC 645
QY 557 TGTGTGGCGGGAACCGCGGAGGAGCTATGCGAGGCGACTCGGAGGCGCCGTGGTGT 616
Db 646 TGTGTGGCGGGAACCGCGGAGGAGCTATGCGAGGCGACTCGGAGGCGCCGTGGTGT 705
QY 617 GCAAGGTGAATGGACACTGCTGACGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 676
Db 706 GCAAGGTGAATGGACACTGCTGACGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 765
QY 677 AGCCCAACCGGCTTGGCATCTACACCCGTGCTACTACTTGGAGCTGATCCACACT 736
Db 766 AGCCCAACCGGCTTGGCATCTACACCCGTGCTACTACTTGGAGCTGATCCACACT 825

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-978-404B-7

Query Match 61.9%; Score 477.2; DB 2; Length 1219;
Best Local Similarity 77.1%; Pred. No. 1.6e-99;
Matches 581; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

```
OY 6 CCTGAGAAAGATCTGCGGGGCTCAGAGGCCCCCAGACAGTGGCCCTGGCAGT 65
DB 260 CCAAGAGGGGCGATGTTGGGGGACAGAGGCACTGGGAACAAGTGGCCCTGGCAGT 319
OY 66 GAGCCTGAGAGTCCAGGCGCCATCTAGTGCATCTTCTGCGGGGCTCCCTGATCCACC 125
DB 320 GAGCCTTGTGCAATGAAACCTACTGAGGATTTCTGCGGGGCTCCCTGATCCACC 379
OY 126 CCAATGGGTGCTGACCGCGCGGCGTGGGACCGGACGTCAAGATCTGGCGCCCT 185
DB 380 ACAGTGGGTGCTCAGCGGCGCACACTGTGTGGACCGACTATTGCTGATCCCAAGAT 439
OY 186 CAGGTGCAACGCGGAGACACACCTCTACTACAGACACGCTGCTGCCGTACAGCAG 245
DB 440 CAGAGTACAGCTTCAAGGCACTACCTATTATACACGACACACCTGGGTGAGCGG 499
OY 246 GATCATGCTGACCCACAGATCTACACCGCCAGATCGGAGCGGACATCGCCCTGGA 305
DB 500 GATCATACACACCGGACATTTATGACCAACGATGGGCGGACATCGCCCTGTTGA 559
OY 306 GCTGAGAGAGCGCGGTGAAGGTCTCCAGCCACGTCACAGCGTACCCCTGCGCCCTC 365
DB 560 GCTCAAGAACCCCTGAAACATTTTCAGGCATGTCCACCCCGCTGCTGCGCTGCTC 619
OY 366 AGAGACCTTCCCCCGGGGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 425
DB 620 AGAGACCTTCCCCCTGAGGAACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 679
OY 426 TGAGCGCTCCCGCCGATTTCTCTGAGGACGTGAAGTCCCATTAATGAAACCA 485
DB 680 TGTGAGCGCTCCCGCCGATTTCTCTGAGGAGGTGCAAGTCTCCGCTGCTGAAACCA 739
OY 486 CATTTGTGACGCAAAATACACCTTGGCGCTTACACGGGAGACGATCCGCGTCCG 545
DB 740 GCTTTGTGACGCAAAATACCAAGGTGTCTACAGGGGAGCAACATCCATTTGTCG 799
OY 546 TGAGACATGCTGTGCGCGGAACACCGGAGGACATCATCCGAGGCGATCCGGAGG 605
DB 800 AGAGCACATGCTGTGCTGCGGAGCAAGACAGATCTCCGAGGCTGCTCGGAGG 859
OY 606 GCGCCTGCTGTCAGAGTGAATGACCTGCTGAGCGGCGGCTGCTGAGCTGGGCGA 665
DB 860 ACCTTGTGCTCAGAGTGAAGGAGTACTGCTGCAAGGAGGTGTGCTGAGTGGGTGA 919
OY 666 GGGCTGTGCCAGCCCAACCGGCTGGGATCTACACCGCTGCTACCTACTTGGACTG 725
DB 920 GGGCTGTGCTGCGCCCAACGAGGCTGGCATCTACACTGGGCTACCTATTAATTG 979
OY 726 GATCCACCATCTATGTCCTCAAAACCGCTGAAGC 759
DB 980 GATCCACCGCTATGTCCTCAAAAGACTTGTGAATC 1013
```

RESULT 12

US-09-016-366A-14
Sequence 14, Application US/09016366A
Patent No. 5955431
GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.
APPLICANT: Huang, Chifu
TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
TITLE OF INVENTION: INHIBITORS
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:

ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,366A
FILING DATE: January 30, 1998
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/037,090
FILING DATE: 05-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7093
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1108 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-016-366A-14

Query Match 60.9%; Score 469.6; DB 2; Length 1108;
Best Local Similarity 77.2%; Pred. No. 8.1e-98;
Matches 571; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

```
OY 17 GAATGCTGGGGGTGAGGAGGCCCCCAGAGCAAGTGGCCCTGGGAGGTGACCTGAAG 76
DB 125 GCATGCTGGGAGAGCATGAGCTTCTGAGAGTGGCCCTGGGAGGTGACCTGAAT 184
OY 77 TCCAGCGGCCCATGATGATGACCTTCTGGGGGGGCTCCCTCATCCACCCCGTGGGTG 136
DB 185 TTAAATTAACTACTGATGATACATTTCTGCGGAGGCTCTCTCAACCCACAGTGGTGC 244
OY 137 TGACCGCGCGGCGTGCCTGGGAGCGGACGTCAGATCTGGCCGCTCAGGCTGCAAC 196
DB 245 TCACTGCGGACACAGTGTGGGAGCGGACATCAAAAGCCACAGCTTCCGCGGTGACG 304
OY 197 TCGGGAGAGCAGCACTCTACTACTACAGACCACTGCTGCGGCTCAGCAGATCATCTGC 256
DB 305 TTCTGAGAGATATATATATATGAGGACCACTCTCTTTGAAACGGAATCGTGTGC 364
OY 257 ACCCAGATGTCACACCGGCCAGATCGGAGGAGCATCGCTGCTGGAGGTGGAGAGG 316
DB 365 ACCCCCATATTACACGCGCGAGGCTGGGAGAGCTGCTGCGGAGCTTGAAGTGC 424
OY 317 CGGTGAAGGTCTCCAGACGTCACACAGGTACCCCTGCGCCCTGCTCAGAGACTTCC 376
DB 425 CTGTGAATGTCTCCACCCATATATCAACCCATATCCCTGCGCCCTGCTGAGAGCTTCC 484
OY 377 CCGCGGGGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 436
DB 485 CCGCTGGGACATCGTGTGGGAGAGGTGAGAGGTGAGAGGTGAGAGGTGAGAGGTGAG 544
OY 437 CACCGCATTTCTCTGAGAGGAGTGAAGGTCCCATATATGAGAAACACATTTGTGAGC 496
DB 545 CACCTCTTATCTCTGAGAGCAAGTGAAGTTCCTCATTTGTGAGAAACAGCTGTGTAC 604
OY 497 CAAATACACCTTGGCGCTACAGGAGAGAGAGAGCTCGCATCTGCTGCTGAGCATGC 556
```


APPLICATION NUMBER: 60/032,354
FILING DATE: 04-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7090
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1031 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-978-404B-1

Query Match 60.1%; Score 463; DB 2; Length 1031;
Best Local Similarity 76.4%; Pred. No. 2,5e-96;
Matches 568; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

14 AAGAATCGTCGGGGGTAGAGAGCCGCCAGAGCAAGTGGCCCTGGCAGGTAGACCTGA 73
106 AAGGCAATGTGGGGGACAGAGGACATGAGCAAGTGGCCCTGGCAGGTAGACCTGC 165
74 GAGTCACAGGCGCCATCTGATGATGACCTTCTGGGGGGCTCCCTCATCCACCCCAAGTGGG 133
166 GAGCAATGACACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 225
134 TGGTACCGCCCGCGGCGTGGGACCGGACGAGTCAAGGATGCTGCGCCCTCAGAGGTGC 193
226 TGGTACCGCGGCGGACAGTGTGGGACCGGATGCTGATGATGATGATGATGATGATGATGATG 285
194 AACTGGGAGGACGACACCTCTACACGAGACGACGACGACGACGACGACGACGACGACGACG 253
286 ACCTCGGAGGAGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 345
254 TGCACCCACAGTTCACACCGCCGATGAGGAGGACATGCGCCCTGCGGAGCTGGAGG 313
346 CACACCGGACCTTCTACATGTCAGAGATGGGGGACAGATGCGCCCTGCGGAGGACCTGA 405
314 ACCCGGTGAAGGTCTCCAGCCACGTCACACGTCACACCTGCGCCCTGCGGAGGACCT 373
406 ACCGTGTGAACATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 465
374 TCCCGCCGGGAGTCCGCTGCTGGTCTGCTGGGCGATGAGCAATGATGAGCGCC 433
466 TCCCTCAGGAGAGTGTGCTGGTGGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 525
434 TCCCAACCGCATTTCTCTGAGAGCAGGTGAAGTCCCATTAATGAAAAACCATTTTGTG 493
526 TCCCGCCGACATTTCTCTGAGAGGAGTGAAGTCCCATTAATGAAAAACCATTTTGTG 585
494 AGCAAAATACCACTTGGCGGCTTACACGAGGAGAGGACGATGCGCCCTGCGGAGGAC 553
586 ACTTGAAGTATACAAAGTCTCATCAGAGTGAACATTTCCACATTTGTCAGATGAGA 645
554 TCCCTGTGCGGGAACACCGGAGGAGTCAATGCGAGGCGACATCCGAGAGGCGCCCTGG 613
646 TCCCTGTGCTGGGATGAAGGACATGATCTCTGCCAGGCGACATCCGAGAGGCGCTTGG 705
614 TGTGAAGGTGAATGAGCAGTGGCTGCGAGGCGGGGCTGACGCTGGGCGAGGGCTGTG 673
706 TGTGAAGGTGAATGAGCAGTGGCTGCGAGGCGGGGCTGACGCTGGGCGAGGGCTGTG 765
674 CCCAGCCCAACCGGCTGATCTACACCGGCTGATCTACCTACTACTACTACTACTACTACTACT 733
766 CACAGCCCAACCGGCTGATCTACACCGGCTGATCTACCTACTACTACTACTACTACTACTACT 825
734 ACTATGTCCCAAAAAGCCGTGA 756
826 ACTATGTCCCAAAAAGCCGTGA 848

RESULT 15
US-09-016-366A-24
Sequence 24, Application US/09016366A
Patent No. 5955431
GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.
APPLICANT: Huang, Chifu
TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
TITLE OF INVENTION: INHIBITORS
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,366A
FILING DATE: January 30, 1998
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/037,090
FILING DATE: 05-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7093
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 1103 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-016-366A-24

Query Match 57.8%; Score 445.6; DB 2; Length 1103;
Best Local Similarity 75.1%; Pred. No. 2.2e-92;
Matches 556; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

17 GAATGTCGGGGGTAGAGAGCCGCCAGAGGACAGTGGCCCTGGCAGGTAGACCTGAGAG 76
102 GATTTGTGGAGAGAGAGAGGCTTGTGAAGTAAAGTGGCCCTGGCAGGTAGACCTGAGAT 161
77 TCCAGGCGCCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 136
162 TTAATTCAGTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 221
137 TGACCGCGCGCGGCGGCGGAGGAGCGGAGCGTCAAGATGTCGCCGCCCTCAGAGGTGCAAC 196
222 TCACTGCGGCGACAGTGTGGGAGTGCACATCAAAAGCCCAAGCTCTTCCGTGTAAGC 281
197 TCGGGAGGACGACCTCTACTACAGAGACACAGTCTGCGGCTGAGCAGAGATCATCTGTC 256
282 TCGGAGAGAGAT 341
257 ACCACAGCTTCTACACCGCCGAGATGAGCGGAGATGCGCCCTGCTGAGCTGAGAGAC 316
342 ACCCCACTACTACACAGTGAAGATGGGGAGAGATGCGCTGCGTGGAGCTTGAAGACC 401
317 CGGTGAAGTCTTCAGGACGATCAGAGGTCACACCTGCGCCCTGCTCAGAGACCTTCC 376

Db	402	CTGTGAATGTCCTCCACCATTCTCCACCACCATTCCTGCCCCCTGCGCTGGAGAACCTTCC	461
Oy	377	CCCCGGGGATCCCGTGGCTGCTGAGTCACTGGCTGGGGCGATGTGGACATGATGAGCGCCCTCC	436
Db	462	CCCTGGGAGCACTTCTTGCTGGGTGATACAGAGCGTGGGGCGCATTTGATGATGACGACGCTCTCC	521
Oy	437	CACGGCCATTTCCTCTGTAGAGAGGTGAAGGTCCCCAATATGAGAAACACATTTGTGAGC	496
Db	522	TGCGACCTTATCTCTGTGAAGCAATGTGAAGGTCCCATTTGGAAAAACAGCGCTGTGATC	581
Oy	497	CAAAATACCACTCTTGGGGCCCTACACGGGAGACGACGTCGCCATCGTCCGTGAGACATGC	556
Db	582	GGAAGTACCACTACGTGGCTCTTACACAGAGATGATGTTCCTCATTTGTCCAGAGATGCGATGC	641
Oy	557	TGTGTGCCGGGAGACACCCGGAGGGACTCATGCGCAGGGCGACTCCGAGAGGGCCCTGTGT	616
Db	642	TGTGTGCTGAAATATACAGACGAGCGACATCTCTGCCAGGAGATCATGSGGGGCCACTGCTCT	701
Oy	617	GCAAGTGATATGGCACTGCGCTGAGGCGGGCGTGTACCTGGGGCGAGAGCGCTGTGCC	676
Db	702	GCAAGTATGAGGTTACCTGCGCTGCAACACAGAGATGGTGCACCTGGGGCGAGAGGCTGGCGAG	761
Oy	677	AGCCCAACGGGCGTGGCATCTACACCCCGTGTCACTACTACTTBSGATGSGATCCACGACT	736
Db	762	AGGGCAATCGTCTTGCGGACTTTTACACCCGGGAGTACTACTCGGACTGGAATTCCACGCT	821
Oy	737	ATGTCCCCAAAAAGCCGGA	756
Db	822	ATGTCCCTCAGCGTTTCC	841

Search completed: January 31, 2003, 05:25:27
Job time : 72 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2003, 05:19:45 ; Search time 2210 Seconds
(without alignments)
5650.100 Million cell updates/sec

Title: US-09-598-982-20
Perfect score: 771
Sequence: 1 gggccctcagaagaat.....cgtcaagcgccgcgcgt 771

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rnd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	676.2	87.7	1030	14	BM919268 AGENCOURT
2	580.2	75.3	904	14	BO720404 AGENCOURT
3	495.6	64.3	502	14	BQ082929 K-EST0144
4	468.4	60.8	736	12	BG542089 602571309
5	444.4	57.6	917	14	BO721078 AGENCOURT
6	422.2	54.8	1000	13	BI413250 602986390

7	420.4	54.5	639	9	AA131142
8	401.4	52.1	811	12	BG697540
9	400	51.9	678	12	BG483591
10	375	48.6	512	12	BF724180
11	351	45.5	693	12	BM911728
12	350.2	45.4	376	12	BF850308
13	349.6	45.3	692	14	BM989945
14	348	45.1	688	14	BQ003361
15	334.2	43.3	371	12	BF359130
16	314.6	40.8	379	12	BF849396
17	309.4	40.1	474	10	AM823937
18	305.8	39.7	525	13	BI775716
19	304.2	39.5	545	13	BI681216
20	295	38.3	558	13	BI790801
21	289.4	37.5	291	12	BF876081
22	281.6	36.5	537	14	BM708916
23	276.8	35.9	613	10	AM152543
24	273.6	35.5	411	12	BE862270
25	264.8	34.3	538	12	BE751979
26	261.4	33.9	620	9	AA049080
27	259	33.6	395	14	T84325
28	258.6	33.5	413	12	BF849397
29	253.6	32.9	377	14	BQ329260
30	252.2	32.7	282	13	BG957512
31	251.2	32.6	380	13	BG983221
32	246	31.9	676	13	BG964077
33	240.6	31.2	350	14	H44196
34	239	31.0	569	13	BI287271
35	233.6	30.3	457	14	W40969
36	232.4	30.1	573	9	AA039008
37	229.6	29.8	572	14	BM686231
38	223.6	29.0	393	12	BF850309
39	223.2	28.9	616	9	AA131322
40	220.8	28.6	540	10	AM001462
41	219	28.4	564	13	BI790523
42	217	28.1	416	12	BF192575
43	217	28.1	453	12	BF192574
44	217	28.1	464	12	BF192577
45	217	28.1	564	9	AI675311

ALIGNMENTS

RESULT 1
BM919268
LOCUS BM919268 1030 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6715837 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5748459
ACCESSION BM919268
VERSION BM919268.1 GI:19369647
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1030)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
<http://image.llnl.gov>
Plate: LHAM12776 Row: f Column: 20
High quality sequence stop: 719.
Location/Qualifiers
1..1030

FEATURES

source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5748499"
/clone_lib="NIH_MGC_120"
/lab_host="DH10B"
/note="Organ: pooled pancreas and spleen; Vector:
pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."

```

```

BASE COUNT      180 a      356 c      320 g      173 t      1 others
ORIGIN

```

```

Query Match      87.7%; Score 676.2; DB 14; Length 1030;
Best Local Similarity 94.5%; Pred. No. 6.9e-133;
Matches 711; Conservative 0; Mismatches 39; Indels 2; Gaps 1;

```

```

OY 17 GAATGCTGGGGGTACAGAGGCCCCCAGAGCAATGGCCCTGGCAGGTGAGCTGAGAG 76
Db 121 GTATGCTGGGGGTACAGAGGCCCCCAGAGCAATGGCCCTGGCAGGTGAGCTGAGAG 180
OY 77 TCACAGGCCCACTGATGATGCACTTCTCGGGGGCTCCCTCATCCACCCAGTGGTGC 136
Db 181 TCCGGAGACCAATGATGATGCACTTCTCGGGGGCTCCCTCATCCACCCAGTGGTGC 240
OY 137 TGACCGCGCGGCGTGGTGGAGCGGACGTCAGAGATCGCCGCCCTCAGGTGCAAC 196
Db 241 TGACCGCGCGGCGTGGTGGAGCGGACGTCAGAGATCGCCGCCCTCAGGTGCAAC 300
OY 197 TGCGGGAGACGACCTCTACTACAGGACGACGCTGCGGTCGACGAGATCATCGTGC 256
Db 301 TGCGGGAGACGACCTCTACTACAGGACGACGCTGCGGTCGACGAGATCATCGTGC 360
OY 257 ACCCAGACCTTCTACACCGCCAGATCGAGCGGACATCGCCCTGCTGGAGCTGAGAGAC 316
Db 361 ACCCAGACCTTCTACATCATCCAGAGCTGAGCGGATATCGCCCTGCTGGAGCTGAGAGAC 420
OY 317 CGGTGAGAGCTCCAGGACGCTCCACAGCGTACCCCTGCGCTCCACAGACCTTCC 376
Db 421 CGGTGAGATCTCCAGGCGCGCTCCACAGCGTACCTCCCTGCGGAGACCTTCC 480
OY 377 CCGCGGGAGTGCCTGCTGGTGCATGCTGGGCGGATGGGACATGATGAGCGGCTCC 436
Db 481 CCGCGGGAGTGCCTGCTGGTGCATGCTGGGCGGATGGGACATGATGAGCGGCTCC 540
OY 437 CACGCGCATTTCTCTGAAGCAGGTGAAGTCCCATATGGAAGAAACCATTTTGTGACG 496
Db 541 CACGCGCATTTCTCTGAAGCAGGTGAAGTCCCATATGGAAGAAACCATTTTGTGACG 600
OY 497 CAATATACCACTTGGCCCTTACACGGGAGAGAGAGCTGCCATGCTCGGTGACGATCG 556
Db 601 CAATATACCACTTGGCCCTTACACGGGAGAGAGAGCTGCCATGCTCGGTGACGATCG 660
OY 557 TGTGTGCGGGGGAACCGCGGAGGACATGCGAGGGGACATCCGAGGGCCCTGGGTGT 616
Db 661 TGTGTGCGGGGGAACCGCGGAGGACATGCGAGGGGACATCCGAGGGCCCTGGGTGT 720
OY 617 GCAAGGTGAATGCGACCTGCTGACGAGCGGGCGTGTGCTGAGTGGGG--CGAGGGCTGTGC 674
Db 721 GCAAGGTGAATGCGACCTGCTGACGAGCGGGCGTGTGCTGAGTGGGGAGAGAGGGGTGTGC 780
OY 675 CCAAGCCCAACGGGCTGGCATCTACACCGGTGTCACTACTACTTGAAGTGGATCCACCA 734
Db 781 CCAAGCCCAACGGGCTGGCATCTACACCGGTGTCACTACTACTTGAAGTGGATCCACCA 840
OY 735 CTATGTCCCAAAAGCGGTGAAGGGCGCCGC 766
Db 841 CTATGTCCCAAAAGCGGTGAAGGGCGCCGC 872

```

```

RESULT 2
BO720404
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
human
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 904)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://Image.llnl.gov
Plate: LLNL3605 row: m column: 01
High quality sequence stop: 587.
FEATURES
source
Location/Qualifiers
1..904
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6196776"
/clone_lib="Lupski_sympathetic_trunk"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/note="Vector: pCMV-Sport6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
5'-TCGACCCAGCGGCTCG-3' and
5'-GACGTGTTAGATGCGGAGCGGCGGCTT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."
BASE COUNT      149 a      320 c      299 g      136 t
ORIGIN
Query Match      75.3%; Score 580.2; DB 14; Length 904;
Best Local Similarity 90.1%; Pred. No. 1.3e-112;
Matches 666; Conservative 0; Mismatches 43; Indels 30; Gaps 3;
OY 17 GAATGCTGGGGGTACAGAGGCCCCCAGAGCAATGGCCCTGGCAGGTGAGCTGAGAG 76
Db 121 GTATGCTGGGGGTACAGAGGCCCCCAGAGCAATGGCCCTGGCAGGTGAGCTGAGAG 157
OY 77 TCACAGGCCCACTGATGATGCACTTCTCGGGGGCTCCCTCATCCACCCAGTGGTGC 136
Db 158 TCCGGAGACCAATGATGATGCACTTCTCGGGGGCTCCCTCATCCACCCAGTGGTGC 217
OY 137 TGACCGCGCGGCGTGGTGGAGCGGACGTCAGAGATCGCCGCCCTCAGGGTGCAC 196
Db 218 TGACCGCGCGGCGTGGTGGAGCGGACGTCAGAGATCGCCGCCCTCAGGGTGCAC 250
OY 197 TGCGGGAGACGACCTCTACTACAGGACGACGCTGCTCGGTGACGAGATCATCGTGC 256
Db 251 TGCGGGAGACGACCTCTACTACAGGACGACGCTGCTCGGTGACGAGATCATCGTGC 310
OY 257 ACCCAGACCTTCTACACCGCCAGATCGAGCGGACATCGCCCTGAGCTGAGAGAGC 316

```


COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Clontech Laboratories, Inc.
cDNA Library Preparation: Clontech Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: L1CML1521 row: b column: 03
High quality sequence stop: 725.

FEATURES

Location/Qualifiers

```

1..736
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:4695458"
    /clone_11b="NIH_MGC_77"
    /lab_host="DH10B (TI phage-resistant)"
    /note="Organ: lung; Vector: pMDR-IIB (Clontech); Site_1: SfiI (ggcgccctgcgccc); Site_2: SfiI (ggccatcatagcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCCGACGCGCGGCACTG-drr(30)BN-3' (where B = A, C, G or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length inserts and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC library."

```

BASE COUNT

ORIGIN	138 a	236 c	242 g	120 t
Query Match	60.8%	Score 468.4:	DB 12;	Length 736;
Best Local Similarity	90.9%;	Pred. No. 5.2e-89;		
Matches 567;	Conservative 0;	Mismatches 46;	Indels 11;	Gaps 6
OY	17	GAATGTGTGGGGGTAGAGAGCCCCAGAGAGAAGTGCCCTGGCAGGTGAGCCTGAGAG	76	
Dd	113	GCATGCTTGGGGGTAGAGAGCCCCAGAGAGAAGTGCCCTGGCAGGTGAGCCTGAGAG	172	
OY	77	TTCACGGCCCATATGATGACATTTCTGCGGGGGGCTCCCTCATCCACCCTCACTGGTGC	136	
Dd	173	TCCACGGCCCATATGAGTAGACTTCTGCGGGGGGCTCCCTCATCCACCCTCACTGGTGC	232	
OY	137	TCACCGCGCGGGCGGTGCGTGGAGACCGGAGACGTAAAGATTTGGCCGCCCTCAGAGGTCAAC	196	
Dd	233	TCACCGCGCACCGACATTCGCGGGAGCCGGAGGTCAAAGATTTGGCCGCCCTCAGAGGTCAAC	292	
OY	197	TTCGGGAGAGACACCTCTACTACAGAGACAGTGGTGGCGGTCACAGATCATCATGTGC	256	
Dd	293	TTCGGGAGAGACACCTCTACTACAGAGACAGTGGTGGCGGTCACAGATCATCATGTGC	352	
OY	257	ATCCACAGTTCTACACCGCCAGATGCGAGCGACATCGCCCTGCTGGAGCTGGAGGAGC	316	
Dd	353	ATCCACAGTTCTACACCGCCAGATGCGAGCGACATCGCCCTGCTGGAGCTGGAGGAGC	412	
OY	317	CGGTGAAGTCTTCACAGCATGTCACACGCGTCAACCTTGCCTGCTCAGAGACCTTCC	376	
Dd	413	CGGTGAAGTCTTCACAGCATGTCACACGCGTCAACCTTGCCTGCTCAGAGACCTTCC	472	
OY	377	CGCC-GGGGATCCGCGTGGGTGCTGCTGGTGGGGGATGGAGACAATGATGAGCCCTC	435	
Dd	473	CGCCGGGGGATCCGCTGCTGGGTGCTGCTGGTGGGGGATGGAGACAATGATGAGCCCTC	532	
OY	436	CCACGCGATTTCCTCTGAAGCAGG-TGAAGTCCCATTAAATGGAATAAACACATTGTGA	494	
Dd	533	CCACGCGATTTCCTCTGAAGCAGGTTGAAGTCCCATTAAATGGAATAAACACATTGTGA	592	
OY	495	CGCAAAATACACCTTGGCGCCT-ACACGGGAGAGG-ACGTCCGCACTGTCGTCGACGAC	552	
Dd	593	CGCAAAATACACCTTGGCGCCTTAACACGGGAGAGCAAGAGTTCGTCGTCGTCGACG	652	
OY	553	ATGCT-----GTGTGCGGGGAGACCGCGGAGGAGACTCATGTCAGAGGCG--ACTCCGAGG	605	

[illegible]

Query Match	54.8%	Score 422.2	DB 13	Length 1000
Query	479	AAAACCAATTTGTGACGAAATATCCACCTTGGCGCCCTACACGAGACGACGTCGCCGA	538	
Db	161	AAAACCAATTTGTGTACGAAATATCCACCTTGGCGCCCTACACGAGACGACGTCGCCGA	240	
Qy	539	TGCTCCGTGACGA-CATGCTGTGTCCCGGAGACACCCGGAGGGACTCATCCAGAGCGAC	597	
Db	241	TGCTCCGTGACGAGATGCTGTGTCCCGGAGACACCCGGAGGGACTCATCCAGAGCGAC	300	
Qy	598	TCCGGAGGGCCCCCTGTGTGTGCAAGGTGAATGGCACCTGGCTGACAGCGGGCGTGTCAAC	657	
Db	301	TCCGGAGGGCCCCCTGTGTGTGCAAGGTGAATGGCACCTGGCTGACAGCGGGCGTGTCAAC	360	
Qy	658	TGGGGCGAGGGCTGTGTGCCCCAGCCCAACCGGCGCTGGCATCTACACCCGTGTACCTACTAC	717	
Db	361	TGGGGCGAGGGCTGTGTGCCCCAGCCCAACCGGCGCTGGCATCTACACCCGTGTACCTACTAC	420	
Qy	718	TTTGACTGGATCCACCATGTGTCCCAAAAAAGCGGTGAAGCGGCC	763	
Db	421	TGGACTGGATCCACCATGTGTCCCAAAAAAGCGGTGTAGTACGAC	466	
RESULT 6				
LOCUS	B1413250	1000 bp	mRNA	linear
DEFINITION	60298639001 NCI_CGAP_Lu33 Mus musculus CDNA clone IMAGE:5142478 5'			EST 14-AUG-2001
ACCESSION	B1413250			
VERSION	B1413250			
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (bases 1 to 1000)			
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM1351 row: c column: 23 High quality sequence start: 32 High quality sequence stop: 849.			
FEATURES				
source				
	1..1000			
	/organism="Mus musculus"			
	/strain="C57BL/6J"			
	/db_xref="taxon:10090"			
	/clone="IMAGE:5142478"			
	/clone_lib="NCI CGAP Lu33"			
	/tissue_type="Pooled lung tumors"			
	/lab_host="DH10B (phage-resistant)"			
	/note="Organ: Lung; Vector: pRTT30-Pac (Pharmacia) with a modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGGCGCCCTGTCTTTTCTTTTCTTTT 3'] . Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pRTT30 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."			
BASE COUNT	202 a	294 c	274 g	229 t
ORIGIN				1 others

	Best Local Similarity	77.2%	Pred. No. 3.2e-79;	
	Matches 564:	Conservative 0;	Mismatches 163;	Indels 4; Gaps 4;
OY	17	GAAATGTCGGGGGTAGAGGCCCCACAGAGCATGTGCGCTGGCAGGTGAGCTTGAGAG	76	
Db	73	GCATGTGGGAGNAGATAGAGCTTCTGAGAGTAAGTAGTGGCCCTGGCAGGTGAGCTGAAT	132	
OY	77	TCCACGGCCCATACTMGATGCACTTCTGCGGGGGCTCCCTCATCCACCAGTGGGTGC	136	
Db	133	TTAAATTAAACTAGATGATATCTTCTGCGAGGCTCTCTCATCCACCACAGTGGGTGC	192	
OY	137	TACCCGCGCGGCGCTGGCTGGGACCGGACGTCAGAGATCTGGCCGCTCAGGGTGCAC	196	
Db	193	TACACGCGGACACACTGTGTGGGACCGCACATCAAAGGCCACAGACTCTTCGGGTGCACG	252	
OY	197	TGCGGAGAGCAGCAGCTTACTACACAGAGACAGCTGCTCCGGTCAGCAGATCATCTGTC	256	
Db	253	TTTCGTGAGCAGATCTATTAATACTAATGSGGACACAGCTCTCTTTTAACCGGATCTGTGTGC	312	
OY	257	ACCCACAGTTCTACACCGCCAGATGGAGCGGACATGCGCTGCTGGAGCTGAGAGAGC	316	
Db	313	ACCCACATATACACAGCGCGGAGGTGGGGCAGACGTTGCCCTGCTGGAGCTTGAGGHC	372	
OY	317	CAGTAAAGTCTCCACGCCACGTCACACGGTACCGCTGCCCTGCTCTCAAGACCTTCC	376	
Db	373	CTGTAAATGTCTCCACCCATATCCACCCATATCCCTGCCCTGCCGTGGAGACTTCC	432	
OY	377	CCCCGGGATGCGCGTGTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGGCGCTCC	436	
Db	433	CCCCTGGGACATCGTCTGTGGGTGAGAGGCTGGGGGCAATTGATATGACGAGCCTCTCC	492	
OY	437	CACCGCATTTCTCTGAGAAGGATGAAGGTCCCATTAATGAAAAACACATTTGTGA-C	495	
Db	493	CACGTCCTATCTCTGGAAGCAAGTGAAGGTTCCTCATTTGTGAAAACAGCCTGTGTACC	552	
OY	496	GCAAAATACCACTTGGCGCTTACACGGGGAAGCAGCTGCCATCTGTCGTAAGCAGATG	555	
Db	553	GGAAATACCAACACGCGGCTCTACACGGGAGATGATTTTCCCATTTGTCCATGATGGCATG	612	
OY	556	CTGTGTGCGGGGACACCCCGGAGGAGTCATGCCAGGGCGACTCGGGAGGCGCCCTGGTG	615	
Db	613	CTGTGTCTCTGGAATATACAGAGAGACTCTGCCAGGGCGATTCAGGGGGCCACTGGTTC	672	
OY	616	TGCAAGGTGAATGACACCTGGCTGCGAGCGGGCGTGTGCA-GCTGGGGCGAGGGCTGTGC	674	
Db	673	TGCAAGTGAAGTGAAGTGAAGTGGCTGCGAGCAGAGTGGTACAGCTGGGGTGAAGGCTGGCC	732	
OY	675	CCAGCCC-AACCGGCTGTGGCATCTACACCCGTGTACC-TACTACTTGGAGCTGATCCAC	732	
Db	733	AACAGCCCAACAAAGCCTGGCATCTACACCCGGGTGACCATCTACTTGAAGTGGATCCAC	792	
OY	733	CACATGTGCC 743		
Db	793	GCTTATGTCC 803		
RESULT 7	AA131142	639 bp	mRNA	linear EST 14-MAY-1997
LOCUS	AA131142			
DEFINITION	Z131B01.t1 Soares-pregnant_uterus_NHPU Homo sapiens cDNA clone IMAGE:503497 5' similar to gp:M30038 ALPHA-TRYPcase PRECURSOR (HUMAN); mRNA sequence.			
ACCESSION	AA131142			
VERSION	AA131142.1	GI:1692796		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 639)			
AUTHORS	Hillier,L., Lemon,G., Becker,M., Bernaldo,M.F., Chipelli,B., Christoe,S., Dietrich,N., Dubugue,T., Favellio,A., Gish,W., Hawkins ,M., Hultman,M., Kucaba,T., Lacy,M., Le,N., Mardis,E., Moore			


```
Db 398 CGGTGAAGGTCTCCAGCGACGTCACAGCGTCACCCCTGCCCTGCAGAGACCTTCC 457
QY 377 CCCCCGGATCCGTCGTCGCTCACTGCTGGGGCGAGTGAGACATGATG 427
Db 458 CCCCCGGATCCGTCGTCGCTCACTGCTGGGGCGAGTGAGACATGATG 508

RESULT 9
BG483591
LOCUS BG483591 678 bp mRNA linear EST 21-MAR-2001
DEFINITION 60250309F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4616715 5',
ACCESSION BG483591
VERSION BG483591.1 GI:13415870
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 678)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs@email.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
plate: LLCM1372 row: a column: 04
High quality sequence stop: 672.
Location/Qualifiers
1. 678
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4616715"
/clone_id="NIH_MGC_77"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: lung; Vector: pBNR-LIB (Clontech); Site: 1;
SfiI (ggccgctggcgc); Site 2: SfiI (ggccatattggc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCCATTATG6C-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGAGCGCGGCGACATG-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

BASE COUNT 129 a 224 c 215 g 110 t
ORIGIN
Query Match 51.9%; Score 400; DB 12; Length 678;
Best Local Similarity 90.8%; Pred. No. 1.4e-74;
Matches 516; Conservative 0; Mismatches 40; Indels 12; Gaps 8;

QY 17 GAATGCTCGGGGTCTAGAGGCCCCCAGAGCAAGTGCCCTGGCAGTGAGCTGAGAG 76
Db 112 GCATCGTTGGGGTTCAGAGAGGCCCCCAGAGCAAGTGCCCTGGCAGTGAGCTGAGAG 171
QY 77 -TTCAGGCGCCATCTGATGACATTTCTGGCGGGGCTCCCTCATTCAC-CCCCAGTGGGT 134
Db 172 CTCGACGACCATATGATGATCTTCTGCGGGGCTCCCTCATTCACATCCCCAGTGGGT 231
QY 135 GCTGACGCGCGCGCGGTGGGAGC-GGACGTCAAGATCTGGCGCGCTCCAGGCGTC 193
Db 232 GCTACCGGACGACGACTGCGTGGGACCTGGACGTCAAGATCTGGTCAAGGCGTC 291
QY 194 AACTGCGGAGACACACTTACTACAGAGACAGCTGCTCGGTCAGCAGATCATCG 253
|||||
```

```
Db 292 AACTGCGGAGACACACTTACTACAGAGACACAGCTGCTCGGTCAGCAGATCATCG 351
QY 254 TGACACCCAGATTCTACACCCGCCAGATCGAGGCGGAGATGCGCTGTGAGCTGGAGG 313
Db 352 TGACACCCAGATTCTACACCCGCCAGATCGAGGCGGAGATGCGCTGTGAGCTGGAGG 411
QY 314 AGCCGCTGAAGTCTCCAGCCAGCTCCACAGCGTCACG-CTGCCCTCGCTCAGAG-AC 371
Db 412 AGCCGCTGAAGTCTCCAGCCAGCTCCACAGCGTCACGCTACGCTGCCCTCGCTCAGAGAC 471
QY 372 CTTCGCCCGCGGAGTGCCTGCTGCTGCTGCTGCTGGGCGATGTGACAAATGATGAGCG 431
Db 472 CTTCGCCCGCGGAGTGCCTGCTGCTGCTGCTGCTGGGCGATGTGACAAATGATGAGCG 531
QY 432 CTTCGCCCGCGGATTTCTCTGGAAGCAGAGTGAAGTCCCATTAATGAGAAACCATTTTG 491
Db 532 GCTCCACGCG-CATTTCCTGGAAGCAGGTAAAGTCCCATTAATGAGAAACCATTTTG 590
QY 492 TGACG--CAAAATACACCTTGCGGCGCTACACGGG---AGACGACGTCCGATGCTCG 545
Db 591 TGACGCGCAAAATTACACCTTGCGGCGCTACACCGGAGAGACAGAGTCCGACTGCTCG 650
QY 546 TGACGACATGCTGTGTGCGCGGGAACACC 573
Db 651 TGACGACATGCTGTGTGCGCGGGAACACC 678

RESULT 10
BF724180 512 bp mRNA linear EST 05-JAN-2001
LOCUS BF724180
DEFINITION bx01h02.y2 Human Iris cDNA (un-normalized, unamplified): BX Homo
sapiens cDNA clone bx01h02 5', mRNA sequence.
ACCESSION BF724180
VERSION BF724180.1 GI:12040089
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 512)
AUTHORS Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
JOURNAL NEIBANK: EST analysis and bioinformatics for ocular genomics
COMMENT Invest. Ophthalmol. Vis. Sci. 41, (2000) In press
Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 01 row: h column: 02
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers
1. 512
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="bx01h02"
/clone_id="bx01h02"
/lab_host="Human Iris cDNA (un-normalized, unamplified):
BX"
/tissue_type="Iris"
/dev_stage="Adult"
/lab_host="EMDH10B"
/note="Organ: Eye; Vector: pCMVSPORT6; Post-mortem iris
tissue was pooled from 10 individuals ranging in age from
4-80 years and RNA was extracted. From this pooled sample
an aliquot of 60ug of total RNA yielded 2.17ug of mRNA. A
directionally cloned cDNA library in the pCMVSPORT6 vector
was constructed at Life Technologies, essentially
following the protocols of the Superscript Plasmid System
full details of which are contained in the manufacturer's
instruction manual (http://www.lifetechn.com/). First
strand synthesis was carried out using a Not I
primer-adaptor [5'-pGACTAGTTTAAATCGGAGCGGCGGCC(T)15-3'
```

1. Not I/Dlunt end inserts were cloned into the Not I/EcoR
V sites in the vector. EST analysis was performed on the
unplified library at the NIH Intramural Sequencing
Center (NISC).
BASE COUNT 83 a 176 c 167 g 84 t 2 others
ORIGIN

Query Match 48.6%; Score 375; DB 12; Length 512;
Best Local Similarity 93.5%; Pred. No. 2.6e-69;
Matches 390; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 17 GAATGTCGGGGGTCAGAGAGCCGCCAGAGCAAGTGGCCCTGCGAGGTGAGCTGAGAG 76
DB 96 GTATGTCGGGGGTCAGAGAGCCGCCAGAGCAAGTGGCCCTGCGAGGTGAGCTGAGAG 155
QY 77 TCCAGGGCCCATCTGATGATGCTTTCGGGGGGCTCCCTCATCCACCCCGAGTGGGTGC 136
DB 156 TCCGGACCATGATGATGATGCTTTCGGGGGGCTCCCTCATCCACCCCGAGTGGGTGC 215
QY 137 TGACCCGCGCGGCTGCGTGGAGCCGACGTCAAGAGATCTGCCGCCCTCAGGTGCAAC 196
DB 216 TGACCCGCGCGGCTGCGTGGAGCCGACGTCAAGAGATCTGCCGCCCTCAGGTGCAAC 275
QY 197 TCGGGAGACACACCTCTACTACAGACACAGCTGCTGCCGTGACAGAGATCATCTGC 256
DB 276 TCGGGAGACACACCTCTACTACAGACACAGCTGCTGCCGTGACAGAGATCATCTGC 335
QY 257 ACCACAGTTTACACCCGACAGATCGGAGGAGCATGCCCTGTGAGTGGAGAGAGC 316
DB 336 ACCACAGTTTACATCATCTCCAGACTGGAGCGGATATGCCCTGTGAGTGGAGAGAGC 395
QY 317 CGGTGAAGTCTCCAGCCACGTCCACAGGTGACCCCTGCCCTCAGAGACCTTCC 376
DB 396 CCGTGAACATCTNACAGCGCGCTNACAGAGTCAATGCTGCCCTGCCCTGAGACCTTCC 455
QY 377 CCCCCGGGATCCGCTGCTGGGTCTACTGCTGGGCGGATGTGACATGATGACGCC 433
DB 456 CCCCCGGGATCCGCTGCTGGGTCTACTGCTGGGCGGATGTGACATGATGATGACGCC 512

RESULT 11
BM991728/c 693 bp mRNA linear EST 17-JUN-2002
LOCUS BM991728
DEFINITION UI-H-DF1-auk-a-06-0-UI.s1 NCI-CGAP-DF1 Homo sapiens cDNA clone
IMAGE:5870477 3', mRNA sequence.
BM991728
BM991728.1 GI:19711117
EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 693)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
NATIONAL Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgaps-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/BLINL at: <http://image.llnl.gov>
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source location/Qualifiers
1..693
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5870477"
/clone_id="NCI-CGAP-DF1"

/tissue_type="Subchondral Bone"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/note="Organ: Bone; Vector: p773-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
NCI-CGAP-DF1 is a normalized cDNA library containing the
following tissue(s): Subchondral Bone. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into p773-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
GTAAAGCTC.
TAG-LIB-UT-H-DF1
TAG-TISSUE=Subchondral bone
TAG_SEQ=GTAAAGCTC"
BASE COUNT 131 a 179 c 241 g 140 t 2 others
ORIGIN

Query Match 45.5%; Score 351; DB 14; Length 693;
Best Local Similarity 96.7%; Pred. No. 3.3e-64;
Matches 357; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 395 GGGTCACCTGGTGGGGGATGTGACATGATGAGCGCTCCACCGCCATTCCTCTGA 454
DB 693 GGGTCACCTGGTGGGGGATGTGACATGATGAGCGCTCCACCGCCATTCCTCTGA 634
QY 455 AGCAGGTGAAGTCCCATTAATGAAAAACACATTTGTGAGCAAAATACACCTTGGCG 514
DB 633 AGCAGGTGAAGTCCCATTAATGAAAAACACATTTGTGAGCAAAATACACCTTGGCG 574
QY 515 CCTACAGGGAGAGACAGCTCCGATCTGCTGACAGACATGCTGTGCGGGAAACCC 574
DB 573 CCTACAGGGAGAGACAGCTCCGATCTGCTGACAGACATGCTGTGCGGGAAACCC 514
QY 575 GGAGGGACTCATGTGCGAGGGGAGCTCCGAGGGGCCCTGGTGTGCAAGTGAATGGACCT 634
DB 513 GGAGGGACTCATGTGCGAGGGGAGCTCCGAGGGGCCCTGGTGTGCAAGTGAATGGACCT 454
QY 635 GGCTGCAAGGGCGGGGTGTGCTGAGTGGGCGAGGGCTGTGCCACACCGGCTTGCA 694
DB 453 GGCTGCAAGGGCGGGGTGTGCTGAGTGGGCGAGGGCTGTGCCACACCGGCTTGCA 394
QY 695 TCTACACCCGTGTCACTACTACTTGTGACTGGAATCACCACCTATGTCCCAAAAGCCGT 754
DB 393 TTTACACCCGTGTCACTACTACTTGTGACTGGAATCACCACCTATGTCCCAAAAGCCGT 334
QY 755 GAAGCGGC 763
DB 333 GAGTCAAGC 325

RESULT 12
BF850308/c 376 bp mRNA linear EST 16-JAN-2001
LOCUS BF850308
DEFINITION CM3-EN0077-181100-489-e06 EN0077 Homo sapiens cDNA, mRNA sequence.
BM850308
BF850308.1 GI:12237470
EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 376)
Dias Neto E., Garcia Correa R., Verjowski-Almeida S., Briones M.R.,
Nagai M.A., da Silva M., Zago M.A., Bordin S., Costa F.F.,
Goldman G.H., Carvalho A.F., Matsukuma A., Bata G.S., Simpson D.H.,
Brunstein A., Oliveira P.S., Bucher P., Jongeneel C.V., O'Hare
M.J., Soares F., Brentani R.R., Reis L.F., de Souza S.J. and

TITLE Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rue Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=CM3&l2=CM3-EN0077-
181100-489-e06&l3=2000-11-18&l4=1)
Seq primer: puc 18 forward
High quality sequence start: 28
High quality sequence stop: 376.
Location/Qualifiers

FEATURES
source
1..376
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0077"
/dev_stage="Adult"
/note="Organ: Lung_normal; Vector: puc18; Site.1: SmaI;
Site.2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 62 a 115 c 121 g 78 t
ORIGIN

Query Match 45.4%; Score 350.2; DB 12; Length 376;
Best Local Similarity 96.5%; Pred. No. 4.1e-64;
Matches 358; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 14 GCGGCGTGGTGGAGCCGAGCATCTGCGCCCTCAGAGGTGCAACTCGCGGA 203
DB 375 CGCGTTATCGCGGAGCCGAGCATCTGCGCCCTCAGAGGTGCAACTCGCGGA 316
QY 204 GCAGACTCTACTACCGAGACCACTCTGCGCGTCAAGAGATCATCTCACCACA 263
DB 315 GCAGACTCTACTACCGAGACCACTCTGCGCGTCAAGAGATCATCTCACCACA 256
QY 264 GTTCTACACCGCCGAGATCGAGCATGCGCTGTGAGAGTGAAGAGCGGTGA 323
DB 255 GTTCTACACCGCCGAGATCGAGCATGCGCTGTGAGAGTGAAGAGCGGTGA 196
QY 324 GGTCTCAGCCAGTCCACAGCGTGCCTGCGCTCAGAGACCTTCCCGCGG 383
DB 195 GGTCTCAGCCAGTCCACAGCGTGCCTGCGCTCAGAGACCTTCCCGCGG 136
QY 384 GATGCGGTCTGCTGCTCATGCTGCTGGGGGAGATGAGCAATATAGCGCTCCACCGC 443
DB 135 GATGCGGTCTGCTGCTCATGCTGCTGGGGGAGATGAGCAATATAGCGCTCCACCGC 76
QY 444 ATTTCCTTGAAGCAGTGAAGTCCCATATGAAAAACCATTTTGAAGCAAAATA 503
DB 75 ATTTCCTTGAAGCAGTGAAGTCCCATATGAAAAACCATTTTGAAGCAAAATA 11
QY 504 CCACCTTGCGC 514
DB 15 GCACATAAGTG 5

RESULT 13
BM989945/c BM989945 692 bp mRNA linear EST 17-JUN-2002
LOCUS
DEFINITION UI-H-D10-ato-m-22-0-UI.s1 NCI_CGAP_D10 Homo sapiens cDNA clone

IMAGE:5862333 3', mRNA sequence.
ACCESSION BM989945
VERSION BM989945.1 GI:19709334
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 692)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@renail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers

FEATURES
source
1..692
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5862333"
/clone_lib="NCI_CGAP_D10"
/tissue_type="Lung Focal Fibrosis"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/note="Organ: Lung; Vector: pRT73-Pac (Pharmacia) with a
modified polylinker; Site.1: EcoR I; Site.2: Not I;
NCI_CGAP_D10 is a cDNA library containing the following
tissue(s): A pool of Lung Focal Fibrosis. The library was
constructed according to Bonaldi, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adapter, digested with Not I, and cloned directionally
into pRT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dfr)18 tail. The sequence tag for this library is
ATACCGCGTC.
TAG_LIB=UI-H-D10
TAG_TISSUE=Lung with fibrosis
TAG_SEQ=ATACCGCGTC"

BASE COUNT 130 a 182 c 242 g 137 t 1 others
ORIGIN

Query Match 45.3%; Score 349.6; DB 14; Length 692;
Best Local Similarity 98.4%; Pred. No. 6.5e-64;
Matches 363; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 395 GGGTCACTGCTGGGGCCATGTGACATGATGAGCGCTCCACCGCATTTCTCTGA 454
DB 692 GGGTCACTGCTGGGGCCATGTGACATGATGAGCGCTCCACCGCATTTCTCTGA 633
QY 455 ACCAGTGAAGTCCCATATGAAAAACCATTTGTAGCAAAATTCACACTTGGG 514
DB 632 ACCAGTGAAGTCCCATATGAAAAACCATTTGTAGCAAAATTCACACTTGGG 574
QY 515 CCTACAGGAGACGACGTCCCATGCTCCGTGAGACATGCTGTGTGCGGGAACACC 574
DB 573 CMTACACGGAGACGACGTCCCATGCTCCGTGAGACATGCTGTGTGCGGGAACACC 514
QY 575 GAGGAGACTCATGCGCAGGCGACTCCGAGGGCCCTGCTGTGCAAGTGAATGGCACT 634
DB 513 GAGGAGACTCATGCGCAGGCGACTCCGAGGGCCCTGCTGTGCAAGTGAATGGCACT 454
QY 635 GCCTGACAGCGGGCGTGTGCACTGGGGGAGGGGCTGTGCCCAAGCCACGGCTGGCA 694

```

Db      453  GGCTGAGGGGGCTGTGTCAGCTGGGGGAGGGCTGTGCCAGCCCAACCGGCTGGGA 394
Oy      695  TCTACACCGCTGTACCTACTGTGAGACTGTGATCCACACTATGTCCCAAAAGCCGT 754
Db      393  TCTACACCGCTGTACCTACTGTGAGACTGTGATCCACACTATGTCCCAAAAGCCGT 334
Oy      755  GAAGCGGCC 763
Db      333  GAGTCAGGC 325

RESULT 14
B0003361c 688 bp mRNA linear EST 26-MAR-2002
LOCUS      UI-H-E11-42d-d-03-0-UI.s1 NCI_CGAP_E11 Homo sapiens CDNA clone
DEFINITION IMAGE:5847122 3', mRNA sequence.
ACCESSION  B0003361
VERSION     B0003361.1 GI:19728261
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 688)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
            CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be found
            through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
            Seq primer: M13 FORWARD
            POLY-A=Yes.

FEATURES
source      1..688
             Location/Qualifiers
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_image:5847122"
             /clone_lib="NCI_CGAP_E11"
             /tissue_type="Chondrosarcoma"
             /dev_stage="Adult"
             /lab_host="DH10B (Life Technologies)"
             /note="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia)
             with a modified polylinker; Site_1: EcoR I; Site_2: Not I;
             NCI_CGAP_E11 is a normalized CDNA library containing the
             following tissue(s): Chondrosarcoma. The library was
             constructed according to Bonaldi, Lennon and Soares,
             Genome Research, 6:791-806, 1996. First strand CDNA
             synthesis was primed with an oligo-dT primer containing a
             Not I site. Double stranded cDNA was ligated to an EcoR I
             adaptor, digested with Not I, and cloned directionally
             into pT7T3-Pac vector. The oligonucleotide used to prime
             the synthesis of first-strand CDNA contains a library tag
             sequence that is located between the Not I site and the
             (dT)18 tail. The sequence tag for this library is
             ACACCTGCAC.
             TAG_LIB=UI-H-E11
             TAG_TISSUE=chondrosarcoma
             TAG_SEQ=ACACTGTGCAC"

BASE COUNT 130 a 176 c 242 g 140 t
ORIGIN
Query Match 45.1%; Score 348; DB 14; Length 688;
Best Local Similarity 97.3%; Pred. No. 1,4e-63;
Matches 354; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Oy 400 ACTGGCTGGGGCGAGTGGACAATGATGAGCGCTCCACGCCCATTTCTCTGAAGCAG 459

```

```

Db      688  ACTGCTGGGGCGAGTGGACAATGATGAGCGCTCCACGCCCATTTCTCTGAAGCAG 629
Oy      460  GTGAGGTGCCCATATATGGAAGAACACATTTGTGACAGCAAAATACACCTTGAGCCCTAC 519
Db      628  GTGAAGGTGCCCATATATGGAAGAACACATTTGTGACAGCAAAATACACCTTGAGCCCTAC 569
Oy      520  ACGGAGACGACGTCCCGCATCTCGGTGACGACATGTGTGTGCGGGGAACACCCGAGG 579
Db      568  ACGGAGACGACGTCCCGCATCTCGGTGACGACATGTGTGTGCGGGGAACACCCGAGG 509
Oy      580  GACTCATGCGAGGGGACGTCCCGAGGGGCCCTGGGTGTGCAAGGTGAATGGACCTGGCGT 639
Db      508  GACTCATGCGAGGGGACGTCTGGAGGGGCCCTGGGTGTGCAAGGTGAATGGACCTGGCGT 449
Oy      640  CAGGCGGGGCTGTGACGTGTGGGGGAGGGCTGTGGCCAGCCCAACCGGCTGGCATCTAC 699
Db      448  CAGGCGGGGCTGTGACGTGTGGGGGAGGGCTGTGGCCAGCCCAACCGGCTGGCATCTAC 389
Oy      700  ACCCGTGTACCTACTACTGTGGACTGTGATCCACACACTATGTCCCAAAAGCGGTGAAGC 759
Db      388  ACCCGTGTACCTACTACTGTGGACTGTGATCCACACACTATGTCCCAAAAGCGGTGAAGC 329
Oy      760  GGCC 763
Db      328  AGGC 325

RESULT 15
Bf359130/c 371 bp mRNA linear EST 22-NOV-2000
LOCUS      QV3-ET0063-090800-291-f02 ET0063 Homo sapiens CDNA, mRNA sequence.
DEFINITION Bf359130
ACCESSION  Bf359130.1 GI:11318202
VERSION     Bf359130.1
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 371)
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
            Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
            20202663
            Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?l=QV3&t=QV3-ET0063-
            090800-291-f02&t3=2000-08-09&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 12
            High quality sequence stop: 369.

FEATURES
source      1..371
             Location/Qualifiers
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_lib="ET0063"
             /dev_stage="Adult"
             /note="Organ: Lung_tumor; Vector: puc18; Site_1: SmaI;
             Site_2: SmaI; A mini-library was made by cloning products

```

derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions.

BASE COUNT 63 a 116 c 119 g 73 t
ORIGIN

Query Match 43.3%; Score 334.2; DB 12; Length 371;
Best Local Similarity 96.3%; Pred. No. 9.8e-61;
Matches 342; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

```

QY 348 CACCTGCCCCCTGCTCAGAGACCTTCCCCCGGGGATGCCGTGCTGCTCACTGCGCTG 407
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 371 CATGCTGCCCCCTGCTCGAGACCTTCCCCCGGGGATGCCGTGCTGCTCACTGCGCTG 312
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 408 GGGCGATGTGGACAATGATGAGCGCTCCACCGCCATTTCCTCTGAAGCAGGTGAAGT 467
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 311 GGGCGATGTGGACAATGATGAGCGCTCCACCGCCATTTCCTCTGAAGCAGGTGAAGT 252
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 468 CCCCATATGGAAGCAACCATTTGTGAGCAAAATACACCTTGCGCGCTACACGGGAGA 527
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 251 CCCCATATGGAAGCAACCATTTGTGAGCAAAATACACCTTGCGCGCTACACGGGAGA 192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 528 CGAGTCCGCATGCTCGTGAAGCATGCTGTGTGCCGGGAACACCGAGGAGACTCATG 587
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 191 CGAGTCCGCATGCTCGTGAAGCATGCTGTGTGCCGGGAACACCGAGGAGACTCATG 132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 588 CCAGGGCGACTCCGGAGGGCCCCCTGCTGTGCAAGGTGAATGGCACCTGGCTGCAGGGGG 647
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 131 CAAGGGCGACTCTGGAGGGCCCCCTGCTGTGCAAGGTGAATGGCACCTGGCTGCAGGGGG 72
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 648 CGTGGTCAGCTGGGGCGAGGGCTGTGCCCAACCGGCTGGCATCTACACC 702
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 71 CGTGGTCAGCTGGGGCGAGGGCTGTGCCCAACCGGCTGGCATCTACACC 17
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Search completed: January 31, 2003, 06:50:06
Job time : 2242 secs

QY 61 CAGGTGAGCCTGAGAGTCCACGGCCCATCTGATGACCTTCTGCGGGGCTCCCTCATC 120
 Db 61 CAGGTGAGCCTGAGAGTCCACGGCCCATCTGATGACCTTCTGCGGGGCTCCCTCATC 120
 QY 121 CACCCCGAGTGGTGTGACCGCGCGGCGTGTGGAGCCGAGCTCAAGATCTGGCC 180
 Db 121 CACCCCGAGTGGTGTGACCGCGCGGCGTGTGGAGCCGAGCTCAAGATCTGGCC 180
 QY 181 GCCCTCAGGGGTGCAACTGCGGGAGAGCAGCTACTACAGAGCCACCTGCTCGCGGTC 240
 Db 181 GCCCTCAGGGGTGCAACTGCGGGAGAGCAGCTACTACAGAGCCACCTGCTCGCGGTC 240
 QY 241 AGCAGATCATCTGTCACCCACAGTTCTACACCGCCGATCGAGCGAGATCGCCCTG 300
 Db 241 AGCAGATCATCTGTCACCCACAGTTCTACACCGCCGATCGAGCGAGATCGCCCTG 300
 QY 301 CTGAGAGTGGAGAGCGGCTGAAGTCTCCAGCCACGTCACACGCGTACCCCTGCCCT 360
 Db 301 CTGAGAGTGGAGAGCGGCTGAAGTCTCCAGCCACGTCACACGCGTACCCCTGCCCT 360
 QY 361 GCCTCAGAGACCTTCCCGCGGGAGTCCGCTGGGTCACTGCTGGGGGATGTGAC 420
 Db 361 GCCTCAGAGACCTTCCCGCGGGAGTCCGCTGGGTCACTGCTGGGGGATGTGAC 420
 QY 421 AATGATGAGCGCTCCACCGCCATTTCTCTGAAGCAGTGAAGTCCCATATGGA 480
 Db 421 AATGATGAGCGCTCCACCGCCATTTCTCTGAAGCAGTGAAGTCCCATATGGA 480
 QY 481 AACCAATTTTGTAGAGCAAAATACCACCTTGCGGCTTACAGCGAGACGACGTCGCCATC 540
 Db 481 AACCAATTTTGTAGAGCAAAATACCACCTTGCGGCTTACAGCGAGACGACGTCGCCATC 540
 QY 541 GTCGTCAGCAGATGCTGTGCGGGGAACACCGGAGGAGCTCATGTCAGCGGACATCC 600
 Db 541 GTCGTCAGCAGATGCTGTGCGGGGAACACCGGAGGAGCTCATGTCAGCGGACATCC 600
 QY 601 GGAGGGCCCTGTGTGTGCAAGTGAATGGACCTGTGTCAGCGCGGCGTGTGAGCTGG 660
 Db 601 GGAGGGCCCTGTGTGTGCAAGTGAATGGACCTGTGTCAGCGCGGCGTGTGAGCTGG 660
 QY 661 GCGGAGGCGCTGTGTCAGCGCCACACCGGCTGTGATCTACACCGGCTGACTACTG 720
 Db 661 GCGGAGGCGCTGTGTCAGCGCCACACCGGCTGTGATCTACACCGGCTGACTACTG 720
 QY 721 GACTGATCTCACCCTATGTCCCAAAAGCGTGAAGCGCGCGCGTGT 771
 Db 721 GACTGATCTCACCCTATGTCCCAAAAGCGTGAAGCGCGCGCGTGT 771

RESULT 3
 AAZ40175
 ID AAZ40175 standard; DNA; 771 BP.
 AC AAZ40175;
 AC AAZ40175;
 DT 18-FEB-2000 (first entry)
 DE Human beta-tryptase coding sequence.
 DE Human beta-tryptase coding sequence.
 KW Beta-tryptase; human; DNA expression construct; protein production;
 KW combinatorial library screening; X ray crystallography; antigen;
 KW antibody generation; ss.
 OS Homo sapiens.
 OS Homo sapiens.
 PN W09960139-A1.
 PD 25-NOV-1999.
 PF 29-OCT-1998; 98WO-US22994.
 PR 15-MAY-1998; 98US-007970.

XX (PROM-) PROMEGA CORP.
 PA Maffitt MA, Niles AL, Haak-Frendscho M;
 PI WPI: 2000-053300/04.
 DR P-PSDB: AAY55011.
 DR New DNA expression construct for production of enzymatically active
 PT recombinant human beta-tryptase -
 PS Disclosure; Page 43-44; 50pp; English.
 XX This sequence encodes the human beta-tryptase. The invention relates to a
 CC DNA expression construct comprising (5' to 3') a promoter linked to a
 CC signal sequence which is linked to a sequence encoding human
 CC beta-tryptase. The DNA construct is useful for transforming host cells to
 CC express, post translationally process and secrete enzymatically active
 CC human tryptase. The method is useful for the production of large amounts
 CC of tryptase with defined specifications. The transformant is useful for
 CC pharmacological studies, combinatorial library screens and X ray
 CC crystallographic studies. The tryptase produced allows for the
 CC development of tryptase agonists and/or antagonists, is useful as an
 CC antigen to generate antihuman tryptase antibodies in various animals,
 CC can be used in screening for compounds which act as tryptase inhibitors,
 CC antagonists, agonists etc. and to assay for the presence of tryptase in
 CC biological or other solutions. Tryptase inhibitors, antagonists, agonists
 CC etc. may be useful as therapeutics. The tryptase does not require any
 CC post-expression or post-purification modifications or manipulations to
 CC initiate tryptase activity and it has enzymatic activity which compares
 CC favourably with cadaveric tryptase. The availability of enzymatically
 CC active tryptase facilitates the large scale screening of combinatorial
 CC libraries for specific tryptase inhibitors as potential therapeutics and
 CC advances the understanding of the biological significance of tryptase in
 CC mast cell mediated diseases. The tryptase can be used to detect low
 CC levels of tryptase.
 XX
 SQ Sequence 771 BP; 149 A; 256 C; 242 G; 124 T; 0 other;

Query Match 99.2%; Score 764.6; DB 21; Length 771;
 Best local similarity 99.5%; Pred. No. 5e-147;
 Matches 767; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GGGCCCTCGAGAAAGATGCTGGGGGTGAGAGGCCCCAGAGAGTGGCCCTG 60
 Db 1 GGGCCCTCGAGAAAGATGCTGGGGGTGAGAGGCCCCAGAGAGTGGCCCTG 60
 QY 61 CAGGTGAGCCTGAGAGTCCACGGCCCATCTGATGACCTTCTGCGGGGCTCCCTCATC 120
 Db 61 CAGGTGAGCCTGAGAGTCCACGGCCCATCTGATGACCTTCTGCGGGGCTCCCTCATC 120
 QY 121 CACCCCGAGTGGTGTGACCGCGCGGCGTGTGGAGCCGAGCTCAAGATCTGGCC 180
 Db 121 CACCCCGAGTGGTGTGACCGCGCGGCGTGTGGAGCCGAGCTCAAGATCTGGCC 180
 QY 181 GCCCTCAGGGGTGCAACTGCGGGAGAGCAGCTACTACAGAGCCACCTGCTCGCGGTC 240
 Db 181 GCCCTCAGGGGTGCAACTGCGGGAGAGCAGCTACTACAGAGCCACCTGCTCGCGGTC 240
 QY 241 AGCAGATCATCTGTCACCCACAGTTCTACACCGCCGATCGAGCGAGATCGCCCTG 300
 Db 241 AGCAGATCATCTGTCACCCACAGTTCTACACCGCCGATCGAGCGAGATCGCCCTG 300
 QY 301 CTGAGAGTGGAGAGCGGCTGAAGTCTCCAGCCACGTCACACGCGTACCCCTGCCCT 360
 Db 301 CTGAGAGTGGAGAGCGGCTGAAGTCTCCAGCCACGTCACACGCGTACCCCTGCCCT 360
 QY 361 GCCTCAGAGACCTTCCCGCGGGAGTCCGCTGGGTCACTGCTGGGGGATGTGAC 420
 Db 361 GCCTCAGAGACCTTCCCGCGGGAGTCCGCTGGGTCACTGCTGGGGGATGTGAC 420
 QY 421 AATGATGAGCGCTCCACCGCCATTTCTCTGAAGCAGGTGAAGTCCCATATGGA 480
 Db 421 AATGATGAGCGCTCCACCGCCATTTCTCTGAAGCAGGTGAAGTCCCATATGGA 480

Db 421 AATGATGAGCCCTCCACCCGCAATTCCTCTGAAAGCAGTGAAGTCCCATATATGGA 480
QY 481 AACCAATTTGAGAGCAAAATACACCTTGCGGCTACAGGGAGAGATCCGCATC 540
Db 481 AACCAATTTGAGAGCAAAATACACCTTGCGGCTACAGGGAGAGATCCGCATC 540
QY 541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGAGGAGATCATGCCAGGGCATCC 600
Db 541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGAGGAGATCATGCCAGGGCATCC 600
QY 601 GGAGGGCCCCCTGTGTGTGCAAGTGAATGGACACTGCTGCAGCGCGGCTGTCACTGG 660
Db 601 GGAGGGCCCCCTGTGTGTGCAAGTGAATGGACACTGCTGCAGCGCGGCTGTCACTGG 660
QY 661 GGCGAGGCTGTGCCGACCCACCGGCTGACATCTACACCCGTCACCTACTCTG 720
Db 661 GGCGAGGCTGTGCCGACCCACCGGCTGACATCTACACCCGTCACCTACTCTG 720
QY 721 GACTGATCCACACACTATGTCCCAAAAAGCGGTGAAGCGCGCGCTGCT 771
Db 721 GACTGATCCACACACTATGTCCCAAAAAGCGGTGAAGCGCGCGCTGCT 771

RESULT 4
AAS20765
ID AAS20765 standard; DNA; 771 BP.
AC AAS20765;
XX
XX 09-APR-2002 (first entry)
DT
XX
XX DNA encoding recombinant human beta-II tryptase.
DE
XX
XX Human: proteolytic tryptase; protease; recombinant beta-II tryptase;
KM enzyme; gene; mutant; ds.
XX
XX Homo sapiens.
OS
XX Synthetic.
OS
XX
FH Key location/Qualifiers
FT CDS 7..756
FT /*tag= a
FT /partial
FT /product= "Recombinant beta-II tryptase"
XX /note= "This sequence lacks a start codon"

MO200198470-A2.
XX
XX 27-DEC-2001.
XX
XX 20-JUN-2001; 2001WO-US19681.
XX
XX 21-JUN-2000; 2000US-0598982.
XX
XX (PROM-) PROMEGA CORP.
XX
XX Maffitt M, Niles AL, Haak-Frendscho M;
XX
XX WPI: 2002-114578/15.
XX
XX P-PSDB; AAU12007.
XX
XX
XX DNA construct for producing enzymatically-inactive proteolytic
PT tryptase, comprises DNA sequence encoding proteolytic tryptase having
PT an active site mutation -
XX
XX
XX Claim 43; Page 75-77; 126pp; English.
XX
XX The present invention relates to recombinant human proteolytic
CC tryptases, active site mutants of these tryptases and the methods for
CC producing these. The method involves the production of a DNA expression
CC construct comprising a promoter operably linked to a secretion signal
CC sequence which is operably linked to a DNA sequence encoding a
CC proteolytic tryptase with an active site mutation (the construct drives

CC expression of a mature proteolytic tryptase that lacks enzymatic activity
CC due to the active site mutation, in hosts transformed to contain the
CC construct). The method is useful for producing enzymatically-active
CC beta-II tryptase. The active site mutants of proteolytic tryptase provide
CC a tool to investigate the structural and functional properties of the
CC protease and its enzymatic activity, and for modeling studies. The
CC enzymatically-active, recombinant proteolytic tryptase produced are
CC useful as an antigen to generate anti-human tryptase antibodies
CC and in drug screening for compounds which act as tryptase inhibitors,
CC antagonists, agonists, etc. The present sequence encodes for recombinant
CC human beta-II tryptase.
XX
XX
XX Sequence 771 BP; 149 A; 256 C; 242 G; 124 T; 0 other;
SQ
Query Match 99.2%; Score 764.6; DB 24; Length 771;
Best Local Similarity 99.5%; Pred. No. 5e-147;
Matches 767; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGGCCCTCGAGAAAAGATCGTGGGGGTGAGAGAGGCCCCAGAGACAGTGGCCCTG 60
Db 1 GGGCCCTCGAGAAAAGATCGTGGGGGTGAGAGAGGCCCCAGAGACAGTGGCCCTG 60
QY 61 CAGGTGACCTGAGAGTCCACGCCCCATCTGATGACATCTTGCGGGGCTCCCTATC 120
Db 61 CAGGTGACCTGAGAGTCCACGCCCCATCTGATGACATCTTGCGGGGCTCCCTATC 120
QY 121 CACCCCGAGTGGGTGCTGACCGCGCGGCTGCGTGGGAGCGGAGTGAAGATCTGGCC 180
Db 121 CACCCCGAGTGGGTGCTGACCGCGCGGCTGCGTGGGAGCGGAGTGAAGATCTGGCC 180
QY 181 GCGCTCAGAGGTGCAACTGCGGGAGAGCACCCTTACTACTACAGAGACAGACTGCTCGGTC 240
Db 181 GCGCTCAGAGGTGCAACTGCGGGAGAGCACCCTTACTACTACAGAGACAGACTGCTCGGTC 240
QY 241 AGCAGATCATCTGTCACCCACAGTTCTACACCGCCAGATGGAAGCGAGATCGCCTG 300
Db 241 AGCAGATCATCTGTCACCCACAGTTCTACACCGCCAGATGGAAGCGAGATCGCCTG 300
QY 301 CTGAGGCTGAGAGAGCGGCTGAAGTCTCAGCCAGCTGCACAGGTCACCTGCCCCCT 360
Db 301 CTGAGGCTGAGAGAGCGGCTGAAGTCTCAGCCAGCTGCACAGGTCACCTGCCCCCT 360
QY 361 GCTTCAGAGACCTTCCCCCGGGGATGCGGTGCTGGGTCACTGGCTGGGGCGATGTGAC 420
Db 361 GCTTCAGAGACCTTCCCCCGGGGATGCGGTGCTGGGTCACTGGGTGGGATGTGAC 420
QY 421 AATGATGAGCGCTCCACCGCCATTTCTCTGAAGCAGTGAAGTCCCATATGGA 480
Db 421 AATGATGAGCGCTCCACCGCCATTTCTCTGAAGCAGTGAAGTCCCATATGGA 480
QY 481 AACCAATTTGAGAGCAAAATACACCTTGCGGCTACAGGGAGAGATCCGCATC 540
Db 481 AACCAATTTGAGAGCAAAATACACCTTGCGGCTACAGGGAGAGATCCGCATC 540
QY 541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGAGGAGATCATGCCAGGGCATCC 600
Db 541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGAGGAGATCATGCCAGGGCATCC 600
QY 601 GGAGGGCCCCCTGTGTGTGCAAGTGAATGGACACTGCTGCAGCGCGGCTGTCACTGG 660
Db 601 GGAGGGCCCCCTGTGTGTGCAAGTGAATGGACACTGCTGCAGCGCGGCTGTCACTGG 660
QY 661 GGCGAGGCTGTGCCGACCCACCGGCTGACATCTACACCCGTCACCTACTCTG 720
Db 661 GGCGAGGCTGTGCCGACCCACCGGCTGACATCTACACCCGTCACCTACTCTG 720
QY 721 GACTGATCCACACACTATGTCCCAAAAAGCGGTGAAGCGCGCGCTGCT 771
Db 721 GACTGATCCACACACTATGTCCCAAAAAGCGGTGAAGCGCGCGCTGCT 771

RESULT 5
AAS20763

```
ID  AAS20763 standard; DNA; 771 BP.
XX
XX  AAS20763;
AC
XX
XX  09-APR-2002 (first entry)
DT
XX  DNA encoding human beta-I tryptase.
DE
XX
XX  Human: proteolytic tryptase; protease; beta-I tryptase;
KW  enzyme; gene; ds.
OS  Homo sapiens.
XX
XX  Key  Location/Qualifiers
FH  7..756
FT  CDS /tag= a
FT  /partial
FT  /product= "Beta-I tryptase"
FT  /note= "This sequence lacks a start codon"
XX
XX  WO200198470-A2.
XX
XX  27-DEC-2001.
PD
XX
XX  20-JUN-2001; 2001WO-US19681.
XX
XX  21-JUN-2000; 2000US-0598982.
XX
XX  (PROM-) PROMEGA CORP.
PA
XX  Maffit M, Niles AL, Haak-Frendscho M;
PI
XX  WPI: 2002-114578/15.
DR  P-PSDB: AAU12006.
XX
XX  DNA construct for producing enzymatically-inactive proteolytic
PT  tryptase, comprises DNA sequence encoding proteolytic tryptase having
PT  an active site mutation -
XX
XX  Disclosure: Page 72-74; 126pp; English.
PS
XX
XX  The present invention relates to recombinant human proteolytic
CC  tryptases, active site mutants of these tryptases and the methods for
CC  producing these. The method involves the production of a DNA expression
CC  construct comprising a promoter operably linked to a secretion signal
CC  sequence which is operably linked to a DNA sequence encoding a
CC  proteolytic tryptase with an active site mutation (the construct drives
CC  expression of a mature proteolytic tryptase that lacks enzymatic activity
CC  due to the active site mutation, in hosts transformed to contain the
CC  construct). The method is useful for producing enzymatically-active
CC  beta-II tryptase. The active site mutants of proteolytic tryptase provide
CC  a tool to investigate the structural and functional properties of the
CC  protease and its enzymatic activity, and for modelling studies. The
CC  enzymatically-active, recombinant proteolytic tryptase produced are
CC  useful as an antigen to generate anti-human tryptase antibodies
CC  and in drug screening for compounds which act as tryptase inhibitors,
CC  antagonists, agonists, etc. The present sequence encodes for human
CC  beta-I tryptase.
XX
XX  Sequence 771 BP; 149 A; 257 C; 241 G; 124 T; 0 other;
SQ
XX
XX  Query Match 99.0%; Score 763; DB 24; Length 771;
XX  Best Local Similarity 99.4%; Pred. No. 1.1e-146;
XX  Matches 766; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 121 CACCCCGAGTGGTGTGACGCGCGGCGTGGGAGCGGAGTCAGGATCTGGCC 180
   |||||
DB 121 CACCCCGAGTGGTGTGACGCGCGGCGTGGGAGCGGAGTCAGGATCTGGCC 180
QY 181 GCCCTCAGGGTGAACCTCGGGAGAGCAGCCTTACTACACAGACAGCTGCTGGCGTTC 240
   |||||
DB 181 GCCCTCAGGGTGAACCTCGGGAGAGCAGCCTTACTACACAGACAGCTGCTGGCGTTC 240
QY 241 AGCAGATATCTGTGTCACCCACAGTTCCTACACCGCCAGATGGGAGCGGACATGCCCTG 300
   |||||
DB 241 AGCAGATATCTGTGTCACCCACAGTTCCTACACCGCCAGATGGGAGCGGACATGCCCTG 300
QY 301 CTGAGCTGAGAGAGCCGCTGAAGGTCTCAGCCACGCTCAGCAGCGTACCTGCCCTCT 360
   |||||
DB 301 CTGAGCTGAGAGAGCCGCTGAAGGTCTCAGCCACGCTCAGCAGCGTACCTGCCCTCT 360
QY 361 GCCTCAGAGACCTTCCCGCGGGGATGCCGTGCTGGTCTACTGGCTGGGCGCATGTGAC 420
   |||||
DB 361 GCCTCAGAGACCTTCCCGCGGGGATGCCGTGCTGGTCTACTGGCTGGGCGCATGTGAC 420
QY 421 AATGATGAGCGCTCCCGACCGCATTTCTCTGAAGCAGGTGAAGTCCCATATGGA 480
   |||||
DB 421 AATGATGAGCGCTCCCGACCGCATTTCTCTGAAGCAGGTGAAGTCCCATATGGA 480
QY 481 AACCATTTTGTGAGCGCAAAATATACACCTTGCGGCTACAGGGAGAGAGCTCCGCATC 540
   |||||
DB 481 AACCATTTTGTGAGCGCAAAATATACACCTTGCGGCTACAGGGAGAGAGCTCCGCATC 540
QY 541 GTCCGTGACGACATCTGTGTGCCGGGAACACCGGAGGAGCTATGCCAGGCGCATTC 600
   |||||
DB 541 GTCCGTGACGACATCTGTGTGCCGGGAACACCGGAGGAGCTATGCCAGGCGCATTC 600
QY 601 GGAGGCGCCCTGTGTGCAAGGTGAATGGACACTGGCTGTCAGAGCGGGGTGTCAGCTGG 660
   |||||
DB 601 GGAGGCGCCCTGTGTGCAAGGTGAATGGACACTGGCTGTCAGAGCGGGGTGTCAGCTGG 660
QY 661 GCGAGAGGCTGTGCCAGCGCAACCGGCTGACATCTACCCGCTGTCACCTACTGTTG 720
   |||||
DB 661 GCGAGAGGCTGTGCCAGCGCAACCGGCTGACATCTACCCGCTGTCACCTACTGTTG 720
QY 721 GACTGGATTCACACTATGTCCCAAAAAACCGGTGAGCGGCGCGTGT 771
   |||||
DB 721 GACTGGATTCACACTATGTCCCAAAAAACCGGTGAGCGGCGCGTGT 771
RESULT 6
AAS20776
ID  AAS20776 standard; DNA; 771 BP.
AC  AAS20776;
XX
XX  09-APR-2002 (first entry)
DT
XX  DNA encoding human beta-II tryptase active site mutant D91A #1.
DE
XX
XX  Human: proteolytic tryptase; protease; recombinant beta-II tryptase;
KW  enzyme; mutant; ds.
XX
XX  Homo sapiens.
OS  Synthetic.
XX
XX  WO200198470-A2.
XX
XX  27-DEC-2001.
PD
XX
XX  20-JUN-2001; 2001WO-US19681.
XX
XX  21-JUN-2000; 2000US-0598982.
XX
XX  (PROM-) PROMEGA CORP.
PA
XX  Maffit M, Niles AL, Haak-Frendscho M;
PI
XX
```


QY 1 GGGCCCTCGAGAAAAGATGTCGGGGGTCAGAGAGCCCCAGAGAGCAAGTGGCCCTGG 60
 |||||||
 Db 1 GGGCCCTCGAGAAAAGATGTCGGGGGTCAGAGAGCCCCAGAGAGCAAGTGGCCCTGG 60
 QY 61 CAGGTGAGCCTGAGAGTCCACGGCCCATCTGATGATGACTTCTGCGGGGGCTCCCTCATC 120
 |||||||
 Db 61 CAGGTGAGCCTGAGAGTCCACGGCCCATCTGATGATGACTTCTGCGGGGGCTCCCTCATC 120
 QY 121 CACCCCGAGTGGGTGCTGACCGCCGGCGGTGCGTGGAGACCGGAGTCAAGATTTGGCC 180
 |||||||
 Db 121 CACCCCGAGTGGGTGCTGACCGCCGGCGGTGCGTGGAGACCGGAGTCAAGATTTGGCC 180
 QY 181 GCGCTCAGAGGTGCAACTGCGGAGAGCAGCCTTACTACAGAGACAGCTGCTGGCGGTC 240
 |||||||
 Db 181 GCGCTCAGAGGTGCAACTGCGGAGAGCAGCCTTACTACAGAGACAGCTGCTGGCGGTC 240
 QY 241 AGCAGGATCATGTCGACCCACAGTTCTACACCGCCAGATCGAGCGACATCGCCCTG 300
 |||||||
 Db 241 AGCAGGATCATGTCGACCCACAGTTCTACACCGCCAGATCGAGCGACATCGCCCTG 300
 QY 301 CTGGAGCTGAGGAGCGGTGAAGTCTCCAGCCAGCTGTCACAGCTGACCCCTGCCCCCT 360
 |||||||
 Db 301 CTGGAGCTGAGGAGCGGTGAAGTCTCCAGCCAGCTGTCACAGCTGACCCCTGCCCCCT 360
 QY 361 GCGCTCAGAGACCTTCCCGCCGGGATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 |||||||
 Db 361 GCGCTCAGAGACCTTCCCGCCGGGATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 QY 421 AATGATGAGCGGCTCCACCGCCATTTCTGTAAGACAGGTGAAGTCCCATTAATGGA 480
 |||||||
 Db 421 AATGATGAGCGGCTCCACCGCCATTTCTGTAAGACAGGTGAAGTCCCATTAATGGA 480
 QY 481 AACCCATTGTAAGGCAAAATACACCTTGCGCTGACGAGGAGAGAGAGTCCGCAATC 540
 |||||||
 Db 481 AACCCATTGTAAGGCAAAATACACCTTGCGCTGACGAGGAGAGAGAGTCCGCAATC 540
 QY 541 GTCCGTGACGACATGCTGTGTCGGGGAACACCCGAGGAGATCTAGCCAGGCGACTCC 600
 |||||||
 Db 541 GTCCGTGACGACATGCTGTGTCGGGGAACACCCGAGGAGATCTAGCCAGGCGACTCC 600
 QY 601 GGAGGGCCCTGCTGTCGAAGGTGAATGGCACCTGCTGACAGGGCGGCTGCTGACTGG 660
 |||||||
 Db 601 GGAGGGCCCTGCTGTCGAAGGTGAATGGCACCTGCTGACAGGGCGGCTGCTGACTGG 660
 QY 661 GCGGAGGCTGTGCGCCACCGACCGGCTGGCATCTACACCCGTGCACTACTGCTTG 720
 |||||||
 Db 661 GCGGAGGCTGTGCGCCACCGACCGGCTGGCATCTACACCCGTGCACTACTGCTTG 720
 QY 721 GACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 771
 |||||||
 Db 721 GACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 771
 RESULT 8
 AAS20777
 ID AAS20777 standard; DNA; 771 BP.
 XX AAS20777;
 AC
 XX 09-APR-2002 (first entry)
 DT
 XX
 DE DNA encoding human beta-II tryptase active site mutant S194A #1.
 KW
 XX Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
 KN enzyme; mutant; ds.
 XX
 XX Homo sapiens.
 OS
 XX Synthetic.
 XX
 PN WO200198470-A2.
 XX
 PD 27-DEC-2001.

XX 20-JUN-2001; 2001WO-US19681.
 PF
 XX
 XX 21-JUN-2000; 2000US-0598982.
 PR
 XX
 XX (PROM-) PROMEGA CORP.
 PA
 XX
 XX Maffei M, Niles AL, Haak-Frendscho M;
 PI
 XX
 XX WPI: 2002-114578/15.
 DR
 XX
 XX P-PSDB: AAU12011.
 PT
 XX
 XX DNA construct for producing enzymatically-inactive proteolytic
 PT tryptase, comprises DNA sequence encoding proteolytic tryptase having
 PT an active site mutation -
 XX
 XX
 PS Claim 7; Page 88-90; 126pp; English.
 CC
 XX
 CC The present invention relates to recombinant human proteolytic
 CC tryptases, active site mutants of these tryptases and the methods for
 CC producing these. The method involves the production of a DNA expression
 CC construct comprising a promoter operably linked to a secretion signal
 CC sequence which is operably linked to a DNA sequence encoding a
 CC proteolytic tryptase with an active site mutation (the construct drives
 CC expression of a mature proteolytic tryptase that lacks enzymatic activity
 CC due to the active site mutation, in hosts transformed to contain the
 CC construct). The method is useful for producing enzymatically-active
 CC beta-II tryptase. The active site mutants of proteolytic tryptase provide
 CC a tool to investigate the structural and functional properties of the
 CC protease and its enzymatic activity, and for modelling studies. The
 CC enzymatically-active, recombinant proteolytic tryptase produced are
 CC useful as an antigen to generate anti-human tryptase antibodies
 CC and in drug screening for compounds which act as tryptase inhibitors,
 CC antagonists, agonists, etc. AAS20775-AAS20790 encode for recombinant
 CC human beta-II tryptase active site mutants.
 XX
 XX
 Sequence 771 BP; 150 A; 255 C; 241 G; 125 T; 0 other;
 50
 Query Match 97.9%; Score 755; DB 24; Length 771;
 Best Local Similarity 98.7%; Pred. No. 4.5e-145;
 Matches 761; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 1 GGGCCCTCGAGAAAAGATGTCGGGGGTCAGAGAGCCCCAGAGAGCAAGTGGCCCTGG 60
 |||||||
 Db 1 GGGCCCTCGAGAAAAGATGTCGGGGGTCAGAGAGCCCCAGAGAGCAAGTGGCCCTGG 60
 QY 61 CAGGTGAGCCTGAGAGTCCACGGCCCATCTGATGATGACTTCTGCGGGGGCTCCCTCATC 120
 |||||||
 Db 61 CAGGTGAGCCTGAGAGTCCACGGCCCATCTGATGATGACTTCTGCGGGGGCTCCCTCATC 120
 QY 121 CACCCCGAGTGGGTGCTGACCGCCGGCGGTGCGTGGAGACCGGAGTCAAGATTTGGCC 180
 |||||||
 Db 121 CACCCCGAGTGGGTGCTGACCGCCGGCGGTGCGTGGAGACCGGAGTCAAGATTTGGCC 180
 QY 181 GCGCTCAGAGGTGCAACTGCGGAGAGCAGCCTTACTACAGAGACAGCTGCTGGCGGTC 240
 |||||||
 Db 181 GCGCTCAGAGGTGCAACTGCGGAGAGCAGCCTTACTACAGAGACAGCTGCTGGCGGTC 240
 QY 241 AGCAGGATCATGTCGACCCACAGTTCTACACCGCCAGATCGAGCGACATCGCCCTG 300
 |||||||
 Db 241 AGCAGGATCATGTCGACCCACAGTTCTACACCGCCAGATCGAGCGACATCGCCCTG 300
 QY 301 CTGGAGCTGAGGAGCGGTGAAGTCTCCAGCCAGCTGTCACAGCTGACCCCTGCCCCCT 360
 |||||||
 Db 301 CTGGAGCTGAGGAGCGGTGAAGTCTCCAGCCAGCTGTCACAGCTGACCCCTGCCCCCT 360
 QY 361 GCGCTCAGAGACCTTCCCGCCGGGATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 |||||||
 Db 361 GCGCTCAGAGACCTTCCCGCCGGGATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 QY 421 AATGATGAGCGGCTCCACCGCCATTTCTGTAAGACAGGTGAAGTCCCATTAATGGA 480
 |||||||
 Db 421 AATGATGAGCGGCTCCACCGCCATTTCTGTAAGACAGGTGAAGTCCCATTAATGGA 480

KW Human; proteolytic trypsinase; protease; recombinant beta-II trypsinase;
 KW enzyme; mutant; ds.
 OS Homo sapiens.
 OS Synthetic.
 XX W0200198470-A2.
 XX PN 27-DEC-2001.
 XX PD 20-JUN-2001; 2001WO-US19681.
 XX PR 21-JUN-2000; 2000US-0598982.
 XX (PROM-) PROMEGA CORP.
 XX Maffit M, Niles AL, Haak-Frendscho M;
 PI WPI; 2002-114578/15.
 DR P-PSDB; AAU12019.
 XX DNA construct for producing enzymatically-inactive proteolytic
 PT trypsinase, comprises DNA sequence encoding proteolytic trypsinase having
 PT an active site mutation -
 PS Claim 7; Page 110-111; 126pp; English.
 XX The present invention relates to recombinant human proteolytic
 CC trypsinases; active site mutants of these trypsinases and the methods for
 CC producing these. The method involves the production of a DNA expression
 CC construct comprising a promoter operably linked to a secretion signal
 CC sequence which is operably linked to a DNA sequence encoding a
 CC proteolytic trypsinase with an active site mutation (the construct drives
 CC expression of a mature proteolytic trypsinase that lacks enzymatic activity
 CC due to the active site mutation, in hosts transformed to contain the
 CC construct). The method is useful for producing enzymatically-inactive
 CC beta-II trypsinase. The active site mutants of proteolytic trypsinase provide
 CC a tool to investigate the structural and functional properties of the
 CC protease and its enzymatic activity, and for modelling studies. The
 CC enzymatically-active, recombinant proteolytic trypsinase produced are
 CC useful as an antigen to generate anti-human trypsinase antibodies
 CC and in drug screening for compounds which act as trypsinase inhibitors,
 CC antagonists, agonists, etc. AAS20775-AAS20790 encode for recombinant
 CC human beta-II trypsinase active site mutants.
 XX
 SQ Sequence 771 BP; 150 A; 256 C; 240 G; 125 T; 0 other;
 Query Match 97.7%; Score 753.4; DB 24; Length 771;
 Best Local Similarity 98.6%; Pred. No. 9.6e-145;
 Matches 760; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 QY 1 GGGCCCCCGAGAAAAAGATGTCGGGGGTCAGAGAGCCCCCAGAGCAAGTGGCCCTGG 60
 DB 1 GGGCCCCCGAGAAAAAGATGTCGGGGGTCAGAGAGCCCCCAGAGCAAGTGGCCCTGG 60
 QY 61 CAGGTGAGCTGAGAGTCCAGGCCCCATCTGATGATCACTTCTGGGGGGCTCCCTCATC 120
 DB 61 CAGGTGAGCTGAGAGTCCAGGCCCCATCTGATGATCACTTCTGGGGGGCTCCCTCATC 120
 QY 121 CACCCCGAGTGGTGTGACCGCCGCGGTGCGTGGAGACCGAGCTCAAGATCTGGCC 180
 DB 121 CACCCCGAGTGGTGTGACCGCCGCGGTGCGTGGAGACCGAGCTCAAGATCTGGCC 180
 QY 181 GCCCTCAGAGGTGCAACTGCGGGAGAGCACTCTACTCTCAAGAGCACTGCTGCCGCTC 240
 DB 181 GCCCTCAGAGGTGCAACTGCGGGAGAGCACTCTACTCTCAAGAGCACTGCTGCCGCTC 240
 QY 241 AGCAGATCATGTCGACCCACAGTTCTACACCGCCGAGATCGGAGACATGCCCGTG 300
 DB 241 AGCAGATCATGTCGACCCACAGTTCTACACCGCCGAGATCGGAGACATGCCCGTG 300
 QY 301 CTGAGCTGAGAGACCGGTGAAGTCTCCAGCCACGTCACACGCTGACCTGCCCTT 360
 DB 301 CTGAGCTGAGAGACCGGTGAAGTCTCCAGCCACGTCACACGCTGACCTGCCCTT 360

DB 301 CTGAGCTGAGAGACCGGTGAAGTCTCCAGCCACGTCACACGCTGACCTGCCCTT 360
 QY 361 GCGTCAGAGACCTTCCCGCGGGGATGCGTGGTGCACGCTGGGGGATGTGAC 420
 DB 361 GCGTCAGAGACCTTCCCGCGGGGATGCGTGGTGCACGCTGGGGGATGTGAC 420
 QY 421 AATGATGAGCGCTCCACCGCCATTTCTGTAAGCAGTGAAGTCCCATTAATGAA 480
 DB 421 AATGATGAGCGCTCCACCGCCATTTCTGTAAGCAGTGAAGTCCCATTAATGAA 480
 QY 481 AACCATTTGTGACGCAAAATACCATTTGCGCTTACAGGGAGACGAGTCCGATC 540
 DB 481 AACCATTTGTGACGCAAAATACCATTTGCGCTTACAGGGAGACGAGTCCGATC 540
 QY 541 GTCCCTGACGACATGCTGTGCGGGGAACCCGAGAGGACTATGCGAGGCGACTCC 600
 DB 541 GTCCCTGACGACATGCTGTGCGGGGAACCCGAGAGGACTATGCGAGGCGACTCC 600
 QY 601 GGAGGGCCCTGTGTGCAAGGTGATGACCTGCTGCAAGGCGGGGTGAGTGTG 660
 DB 601 GGAGGGCCCTGTGTGCAAGGTGATGACCTGCTGCAAGGCGGGGTGAGTGTG 660
 QY 661 GCGGAGGCTGTGCGCCAGCCCAACCGGCTGCGATCTACACCCGTGCTACTACTTG 720
 DB 661 GCGGAGGCTGTGCGCCAGCCCAACCGGCTGCGATCTACACCCGTGCTACTACTTG 720
 QY 721 GACTGATTCACACCTATGTCCCAAAAAGCCGTAAACCGCGCCGCTGT 771
 DB 721 GACTGATTCACACCTATGTCCCAAAAAGCCGTAAACCGCGCCGCTGT 771
 RESULT 11
 ID AAS20786
 XX AAS20786 standard; DNA; 771 BP.
 AC AAS20786;
 XX 09-APR-2002 (first entry)
 DT
 XX DNA encoding human beta-II trypsinase active site mutant S194A #6.
 DE Human; proteolytic trypsinase; protease; recombinant beta-II trypsinase;
 KW enzyme; mutant; ds.
 KW
 XX Homo sapiens.
 OS Synthetic.
 OS W0200198470-A2.
 XX PN 27-DEC-2001.
 XX PD 20-JUN-2001; 2001WO-US19681.
 XX PR 21-JUN-2000; 2000US-0598982.
 XX (PROM-) PROMEGA CORP.
 XX Maffit M, Niles AL, Haak-Frendscho M;
 PI WPI; 2002-114578/15.
 DR P-PSDB; AAU12020.
 XX DNA construct for producing enzymatically-inactive proteolytic
 PT trypsinase, comprises DNA sequence encoding proteolytic trypsinase having
 PT an active site mutation -
 PS Claim 7; Page 113-114; 126pp; English.
 XX The present invention relates to recombinant human proteolytic
 CC trypsinases; active site mutants of these trypsinases and the methods for
 CC producing these. The method involves the production of a DNA expression
 CC construct comprising a promoter operably linked to a secretion signal
 CC sequence which is operably linked to a DNA sequence encoding a

CC proteolytic trypsin with an active site mutation (the construct drives
CC expression of a mature proteolytic trypsin that lacks enzymatic activity
CC due to the active site mutation, in hosts transformed to contain the
CC construct). The method is useful for producing enzymatically-active
CC beta-II trypsin. The active site mutants of proteolytic trypsin provide
CC a tool to investigate the structural and functional properties of the
CC protease and its enzymatic activity, and for modelling studies. The
CC enzymatically-active, recombinant proteolytic trypsin produced are
CC useful as an antigen to generate anti-human trypsin antibodies
CC and in drug screening for compounds which act as trypsin inhibitors,
CC antagonists, agonists, etc. AAS20775-AAS20790 encode for recombinant
CC human beta-II trypsin active site mutants.
CC
XX
XX
SQ Sequence 771 BP; 152 A; 256 C; 240 G; 123 T; 0 other;

Query Match 97.7%; Score 753.4; DB 24; Length 771;
Best Local Similarity 98.6%; Pred. No. 9.6e-141;
Matches 760; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 1 GGGCCCTCGAGAAAAAGATCTGGGGGTCAAGAGGCCGCCAGAGCAAGTGGCCCTGG 60
DB 1 GGGCCCTCGAGAAAAAGATCTGGGGGTCAAGAGGCCGCCAGAGCAAGTGGCCCTGG 60
OY 61 CAGGTGAGCCTGAGAGTCCAGCGCCCATCTGATGACCTTCTGGGGGCTCCCTCATC 120
DB 61 CAGGTGAGCCTGAGAGTCCAGCGCCCATCTGATGACCTTCTGGGGGCTCCCTCATC 120
OY 121 CACCCCGAGTGGGTCTGACCGCGCGGGGTGAGGAGCCGGAGCTCAAGATCTGGGC 180
DB 121 CACCCCGAGTGGGTCTGACCGCGCGGGGTGAGGAGCCGGAGCTCAAGATCTGGGC 180
OY 181 GCCCTCAGAGGTCAACCTGGGGAGAGCAGCCTCTACTACAGAGCAGCTGTCGCGGTC 240
DB 181 GCCCTCAGAGGTCAACCTGGGGAGAGCAGCCTCTACTACAGAGCAGCTGTCGCGGTC 240
OY 241 AGCAGATCATGTGTGACCCAGATTTCTACACCGCCAGATGGAGCGACATCGCCCTG 300
DB 241 AGCAGATCATGTGTGACCCAGATTTCTACACCGCCAGATGGAGCGACATCGCCCTG 300
OY 301 CTGGAGCTGAGAGGACCGGTGAAGTCTCAGCCAGCTCAACAGGTCACCTGGCCCT 360
DB 301 CTGGAGCTGAGAGGACCGGTGAAGTCTCAGCCAGCTCAACAGGTCACCTGGCCCT 360
OY 361 GCCTCAGAGACCTTCCCGCGGGGATGCGGTCTGAGTCTGAGTGGGGCGATGTGAC 420
DB 361 GCCTCAGAGACCTTCCCGCGGGGATGCGGTCTGAGTCTGAGTGGGGCGATGTGAC 420
OY 421 AATGATGAGCGCTCCACCGCCATTTCTCTGAAGCAGGTAAAGTCCCATTAATGAA 480
DB 421 AATGATGAGCGCTCCACCGCCATTTCTCTGAAGCAGGTAAAGTCCCATTAATGAA 480
OY 481 AACCAATTTGGAGCAAAATATACACCTTGGCGCTTACAGGGAGACAGCTCGGCATC 540
DB 481 AACCAATTTGGAGCAAAATATACACCTTGGCGCTTACAGGGAGACAGCTCGGCATC 540
OY 541 GTCCGTGACGACATCTGTGTGCGGGAAACACCGGAGGACTTATGCCAGGCGACTTC 600
DB 541 GTCCGTGACGACATCTGTGTGCGGGAAACACCGGAGGACTTATGCCAGGCGACTTC 600
OY 601 GAGAGGCGCTGTGTGCAAGGTGAATGGACCTGTGTCAGCGCGGGCTGTGACCTGG 660
DB 601 GAGAGGCGCTGTGTGCAAGGTGAATGGACCTGTGTCAGCGCGGGCTGTGACCTGG 660
OY 661 GCGGAGGCGTGGCCAGCGCAACCGGCTGTGATCAACCGGTGTACCTACTACTTG 720
DB 661 GCGGAGGCGTGGCCAGCGCAACCGGCTGTGATCAACCGGTGTACCTACTACTTG 720
OY 721 GACTGATTCACACCTATGTCCCAAAAAAGCGGTGAAGCGCGCGCTGT 771
DB 721 GACTGATTCACACCTATGTCCCAAAAAAGCGGTGAAGCGCGCGCTGT 771

RESULT 12

AAS20779
ID AAS20779 standard: DNA: 735 BP.
XX
XX AAS20779;
AC 09-APR-2002 (first entry)
DT
XX
XX DNA encoding human beta-II trypsin active site mutant H44A #2.
DE
XX
XX Human: proteolytic trypsin; protease; recombinant beta-II trypsin;
KW enzyme; mutant; ds.
KW
XX
XX Homo sapiens.
OS
XX Synthetic.
PN WO200198470-A2.
PD 27-DEC-2001.
XX
XX 20-JUN-2001; 2001WO-US19681.
PF
XX
XX 21-JUN-2000; 2000US-0598982.
PR
XX
XX (PROM-) PROMEGA CORP.
PA
XX
XX Maffit M, Niles AL, Haak-Frendscho M;
PI
XX
XX WPI: 2002-114578/15.
DR P-PSDB; AMU12013.
DR
XX
XX DNA construct for producing enzymatically-inactive proteolytic
PT trypsin, comprises DNA sequence encoding proteolytic trypsin having
PT an active site mutation -
PT
XX
XX
PS Example 1c: Page 94-95; 126pp; English.
XX
XX The present invention relates to recombinant human proteolytic
CC trypsin, active site mutants of these trypsin and the methods for
CC producing these. The method involves the production of a DNA expression
CC construct comprising a promoter operably linked to a secretion signal
CC sequence which is operably linked to a DNA sequence encoding a
CC proteolytic trypsin with an active site mutation (the construct drives
CC expression of a mature proteolytic trypsin that lacks enzymatic activity
CC due to the active site mutation, in hosts transformed to contain the
CC construct). The method is useful for producing enzymatically-active
CC beta-II trypsin. The active site mutants of proteolytic trypsin provide
CC a tool to investigate the structural and functional properties of the
CC protease and its enzymatic activity, and for modelling studies. The
CC enzymatically-active, recombinant proteolytic trypsin produced are
CC useful as an antigen to generate anti-human trypsin antibodies
CC and in drug screening for compounds which act as trypsin inhibitors,
CC antagonists, agonists, etc. AAS20775-AAS20790 encode for recombinant
CC human beta-II trypsin active site mutants.
CC
XX
SQ Sequence 735 BP; 139 A; 245 C; 231 G; 120 T; 0 other;
XX

Query Match 95.3%; Score 735; DB 24; Length 735;
Best Local Similarity 100.0%; Pred. No. 5.4e-141;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 ATCGTGGGGGTCAAGAGGCCGCCAGAGCAAGTGGCCCTGAGAGCTGAGACTG 78
DB 1 ATCGTGGGGGTCAAGAGGCCGCCAGAGCAAGTGGCCCTGAGAGCTGAGACTG 78
OY 79 CAGGCGCCATATGATGATCATTTCTGGGGGCTCCCTCATCCACCCCATGGTGTCTG 138
DB 61 CAGGCGCCATATGATGATCATTTCTGGGGGCTCCCTCATCCACCCCATGGTGTCTG 120
OY 139 ACCGCGGGGGTGGGTGGAGCCGAGGCAAGGATCTGGGGCGCCCTGAGGTGCACTG 198
DB 121 ACCGCGGGGGTGGGTGGAGCCGAGGCAAGGATCTGGGGCGCCCTGAGGTGCACTG 180
OY 199 CGGAGCAGACCTTCTACTACAGAGCAGCTGTGCGGTGACAGAGATCATGTGAC 258


```
Db 181 CGGAGCAGCAGCTTACTACTACAGGACGAGCTGCTGCCGTGACAGATCATCGTGAC 240
Oy 259 CCACAGTCTTACACCGCCAGATGAGGACGATCCGCTGCTGAGAGTGGAGAGCCG 318
Db 241 CCACAGTCTTACACCGCCAGATGAGGACGATCCGCTGCTGAGAGTGGAGAGCCG 300
Oy 319 GTGAGGTCTCCAGGACGATGACGATGACGATGACGATGACGATGACGATGACGAT 378
Db 301 GTGAGGTCTCCAGGACGATGACGATGACGATGACGATGACGATGACGATGACGAT 360
Oy 379 CCGGGGATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 438
Db 361 CCGGGGATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Oy 439 CCGGATTTTCTTACAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 498
Db 421 CCGGATTTTCTTACAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 480
Oy 499 AATATACCACTTGGGCTTACAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 558
Db 481 AATATACCACTTGGGCTTACAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 540
Oy 559 TGTGCGGGAGACACCGGAGGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 618
Db 541 TGTGCGGGAGACACCGGAGGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 600
Oy 619 AAGGTGAATGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 678
Db 601 AAGGTGAATGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Oy 679 CCCAACCGGCTGAGATCTACACCGGATGAGGATGAGGATGAGGATGAGGATGAG 738
Db 661 CCCAACCGGCTGAGATCTACACCGGATGAGGATGAGGATGAGGATGAGGATGAG 720
Oy 739 GTCCCAAAAAAGCG 753
Db 721 GTCCCAAAAAAGCG 735

RESULT 13
AAS20787
ID AAS20787 standard; DNA; 735 BP.
AC AAS20787;
XX
DT 09-APR-2002 (first entry)
DE DNA encoding human beta-II tryptase active site mutant H44A #4.
XX
KW Human; proteolytic trypsin; protease; recombinant beta-II trypsin;
KW enzyme; mutant; ds.
OS Homo sapiens.
OS Synthetic.
XX
PN WO200198470-A2.
XX
PD 27-DEC-2001.
PF 20-JUN-2001; 2001WO-US19681.
XX
PR 21-JUN-2000; 2000US-0598982.
XX
PA (PROM-) PROMEGA CORP.
XX
PI Maffit M, Niles AL, Haak-Frendscho M;
XX
DR WPI: 2002-114578/15.
XX
P-PSDB: AAU12021.
PT DNA construct for producing enzymatically-inactive proteolytic
trypsin, comprises DNA sequence encoding proteolytic trypsin having
```

```
PT an active site mutation -
XX
PS Example 1c; page 115-117; 126pp; English.
CC The present invention relates to recombinant human proteolytic
CC trypsin, active site mutants of these trypsin and the methods for
CC producing these. The method involves the production of a DNA expression
CC construct comprising a promoter operably linked to a secretion signal
CC sequence which is operably linked to a DNA sequence encoding a
CC proteolytic trypsin with an active site mutation (the construct drives
CC expression of a mature proteolytic trypsin that lacks enzymatic activity
CC due to the active site mutation, in hosts transformed to contain the
CC construct). The method is useful for producing enzymatically-inactive
CC beta-II trypsin. The active site mutants of proteolytic trypsin provide
CC a tool to investigate the structural and functional properties of the
CC protease and its enzymatic activity, and for modelling studies. The
CC enzymatically-inactive, recombinant proteolytic trypsin produced are
CC useful as an antigen to generate anti-human trypsin antibodies
CC and in drug screening for compounds which act as trypsin inhibitors,
CC antagonists, agonists, etc. AAS20775-AAS20790 encode for recombinant
CC human beta-II trypsin active site mutants.
SQ Sequence 735 BP; 139 A; 246 C; 230 G; 120 T; 0 other;
Query Match 95.1%; Score 733.4; DB 24; Length 735;
Best Local Similarity 99.9%; Pred. No. 1,1e-140;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 19 ATGTCGGGGGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 78
Db 1 ATGTCGGGGGTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 60
Oy 79 CAGCGGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 138
Db 61 CAGCGGCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120
Oy 139 ACCGCGCGGCGGTGCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 198
Db 121 ACCGCGCGGCGGTGCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
Oy 199 CCGGAGCAGACCTTACTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 258
Db 181 CCGGAGCAGACCTTACTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
Oy 259 CCACAGTCTTACACCGCCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 318
Db 241 CCACAGTCTTACACCGCCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
Oy 319 GTGAGGTCTCCAGGACGATGACGATGACGATGACGATGACGATGACGATGACGAT 378
Db 301 GTGAGGTCTCCAGGACGATGACGATGACGATGACGATGACGATGACGATGACGAT 360
Oy 379 CCGGGGATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 438
Db 361 CCGGGGATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Oy 439 CCGGATTTTCTTACAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 498
Db 421 CCGGATTTTCTTACAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 480
Oy 499 AATATACCACTTGGGCTTACAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 558
Db 481 AATATACCACTTGGGCTTACAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 540
Oy 559 TGTGCGGGAGACACCGGAGGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 618
Db 541 TGTGCGGGAGACACCGGAGGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 600
Oy 619 AAGGTGAATGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 678
Db 601 AAGGTGAATGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Oy 679 CCCAACCGGCTGAGATCTACACCGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 738
```

Db 661 CCAACCGCGCTGGCATCTACACCGCTGTCACTACTGATGATGATCACCACCTAT 720
 QY 739 GTCCCAAAAAGCG 753
 ||||||||||||||||
 Db 721 GTCCCAAAAAGCG 735

RESULT 14
 AAV44330
 ID AAV44330 standard; DNA; 1128 BP.

AC AAV44330;

DT 24-NOV-1998 (first entry)

DE Human mast cell tryptase II/beta nucleic acid sequence.

XX Mast cell tryptase II/beta; human; MCP-7; mast cell protease 7;
 KW blood clot; anticoagulant; myocardial infarction; reocclusion;
 KW thromboembolism; cerebral embolism; thrombosis; therapy; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 2..826

FT CDS /*tag= a

PN WO9824886-A1.

PD 11-JUN-1998.

XX 25-NOV-1997; 97WO-US21620.

XX 04-DEC-1996; 96US-0032354.

XX (BGHM) BRIGHAM & WOMENS HOSPITAL.

XX Stevens RL;

XX WPI: 1998-333308/29.

XX P-PSDB; AAM64240.

XX New compositions containing tryptase-7, e.g. mouse mast cell
 protease-7 - are used to treat clot formation in e.g. myocardial
 infarction, reocclusion following angioplasty or pulmonary
 thromboembolism

XX Disclousure; Page 66; 92pp; English.

XX This nucleotide sequence includes a coding region for human
 CC mast cell tryptase II/beta (see AAM64240). The invention provides:
 CC compositions comprising an isolated tryptase-7 that may include
 CC chimeric proteins that contain (a) a human tryptase for all but
 CC the active site region and (b) the substrate-binding pocket of
 CC mouse tryptase-7 or its homologues (see AAM64233-39); a method for
 CC treating a blood clot by administering a nucleic acid molecule that
 CC codes for a tryptase-7, or an expression product, to decrease
 CC fibrinogen activity; a nucleic acid encoding a serine protease
 CC (SP); and a method of producing a mature SP by expressing the
 CC inactive zymogen in a host cell, and cleaving the enterokinase
 CC susceptibility domain. The tryptase-7 polypeptides can be used to
 CC treat disorders mediated by undesirable thrombus clot formation
 CC such as myocardial infarction and reocclusion following angioplasty
 CC of blood clots associated with pulmonary thromboembolism, deep vein
 CC thrombosis, cerebral embolism, renal vein and peripheral arterial
 CC thrombosis. They are also useful for all surgical procedures that
 CC require decreased blood clots.

XX Sequence 1128 BP; 190 A; 409 C; 329 G; 200 T; 0 other;

Query Match

Best Local Similarity 94.8%; Score 731; DB 19; Length 1128;
 98.7%; Pred. No. 3.6e-140;

Matches 737; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 17 GAATGTCGGGGGTGAGAGAGCCGCCAGAGCAAGTGGCCCTGGAGGTGAGCCTGAGAG 76
 |||||
 Db 87 GCATGCTTGGGGGTGAGAGAGCCGCCAGAGCAAGTGGCCCTGGAGGTGAGCCTGAGAG 146
 QY 77 TCCACGGCCCATCTGATGCACTTCTGGGGGCTCCCTCATCCACCCCACTGGGTGC 136
 |||||
 Db 147 TCACGGCCCATCTGATGCACTTCTGGGGGCTCCCTCATCCACCCCACTGGGTGC 206
 QY 137 TGACCGCGCGCGCGTGGGAGCGGACGTCAGAGATCTGGCGCCCTCAGGGGTGAC 196
 |||||
 Db 207 TGACCGCGAGCGACTGCTGGGACCGGAGCTCAAGAGATCTGGCGCCCTCAGGGGTGAC 266
 QY 197 TCGGAGACAGACAGCCTCTACTACAGAGACAGCTCTCCGCTCAGAGATCATCTGTC 256
 |||||
 Db 267 TCGGAGACAGACAGCCTCTACTACAGAGACAGCTCTCCGCTCAGAGATCATCTGTC 326
 QY 257 ACCCAGACTTCTACACCGGCCAGATCGAGCGGACATCGCCCTGCTGAGAGAGAGC 316
 |||||
 Db 327 ACCCAGACTTCTACACCGGCCAGATCGAGCGGACATCGCCCTGCTGAGAGAGAGC 386
 QY 317 CGGTGAAGGTCTCCAGCCAGCCAGCTCAGAGCTGACCCCTGCTCAGAGACTTTC 376
 |||||
 Db 387 CGGTGAAGGTCTCCAGCCAGCTCAGAGCTGACCCCTGCTCAGAGACTTTC 446
 QY 377 CCCCAGGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436
 |||||
 Db 447 CCCCAGGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 506
 QY 437 CACGCGCATTTCTCTGAAAGCAGGTGAAGTCCCATATGTAAGAAACACATTTGTGAC 496
 |||||
 Db 507 CACGCGCATTTCTCTGAAAGCAGGTGAAGTCCCATATGTAAGAAACACATTTGTGAC 556
 QY 497 CAAATPACCACTTGGCGCTACAGGGAGAGAGAGCTCCGATCGTCCGTGACAGATGC 556
 |||||
 Db 567 CAAATPACCACTTGGCGCTACAGGGAGAGAGAGCTCCGATCGTCCGTGACAGATGC 626
 QY 557 TGTGTGCGGGAACACCGGAGGAGACTGACAGGCGACTCCGAGAGGCCCTGTGTGT 616
 |||||
 Db 627 TGTGTGCGGGAACACCGGAGGAGACTGACAGGCGACTCCGAGAGGCCCTGTGTGT 686
 QY 617 GCAAGGTGATGACACCTGCTGTCAGGCGGCGGTGTCAGCTGGGGCGAGGGCTGTGCC 676
 |||||
 Db 687 GCAAGGTGATGACACCTGCTGTCAGGCGGCGGTGTCAGCTGGGGCGAGGGCTGTGCC 746
 QY 677 AGCCCAACCGGCTGGCATCTACACCCGTCACCTACTGACCTGGAGTGCACCACT 736
 |||||
 Db 747 AGCCCAACCGGCTGGCATCTACACCCGTCACCTACTGACCTGGAGTGCACCACT 806
 QY 737 ATGTCCCAAAAAGCGGTGAAGCGGCC 763
 |||||
 Db 807 ATGTCCCAAAAAGCGGTGAAGCGGCC 833

RESULT 15

AAV42712
 ID AAV42712 standard; CDNA; 1128 BP.

XX AAV42712;

XX 27-OCT-1998 (first entry)

XX Human mast cell tryptase II/beta encoding cDNA.

XX Mast cell protease; MCP; mouse; inhibitor; peptide substrate; asthma;
 KW tryptase-6 protein; inflammatory disorder; allergic rhinitis; urticaria;
 KW antioedema; eczematous dermatitis; atopic dermatitis; anaphylaxis;
 KW hyperproliferative skin disease; peptic ulcer; hypersensitivity;
 KW inflammatory skin condition; human; mast cell tryptase II/beta; ss.

XX Homo sapiens.

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 31, 2003, 04:47:55 ; Search time 2796 Seconds
(without alignments)
8025.130 Million cell updates/sec

Title: US-09-598-982-20

Perfect score: 771

Sequence: 1 gggcccccctcagaagaat.....cgtgaagcggcgcgcgtcgt 771

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_da:*
2: gb_bt:*
3: gb_la:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	771	100.0	771	6	AX347853
2	769.4	99.8	771	6	AX347869
3	764.6	99.2	771	6	ARI65112
4	764.6	99.2	771	6	AX347841
5	763	99.0	771	6	AX347838
6	761.4	98.8	771	6	AX347855
7	759.8	98.5	771	6	AX347871
8	755	97.9	771	6	AX347857
9	755	97.9	771	6	AX347859
10	753.4	97.7	771	6	AX347873
11	753.4	97.7	771	6	AX347875
12	735	95.3	735	6	AX347861
13	733.4	95.1	735	6	AX347877
14	731	94.8	1128	6	AR080461
15	731	94.8	1128	6	HUMTRX2A
16	731	94.8	1137	6	AR080460
17	731	94.8	1137	9	HUMTRX1A
18	731	94.8	1143	9	HUMBRYP
19	731	94.8	1145	9	S5551
20	731	94.8	1194	9	BC029356
21	728.6	94.5	735	6	ARI65109
22	728.6	94.5	735	6	AX347843
23	727	94.3	735	6	AX347834
24	726.2	94.2	1081	6	AR080462
25	726.2	94.2	1081	6	AX329646
26	726.2	94.2	1081	6	AX330205
27	726.2	94.2	1081	6	AX334644
28	726.2	94.2	1081	6	AX375047
29	726.2	94.2	1081	9	HUMTRX3A
30	725.4	94.1	735	6	AX347863
31	723.8	93.9	735	6	AX347879
32	719	93.3	735	6	AX347865
33	719	93.3	735	6	AX347867
34	719	93.3	1221	6	AX014344
35	718.2	93.2	2662	6	AX338508
36	717.4	93.0	735	6	AX347881
37	717.4	93.0	735	6	AX347883
38	692.6	89.8	884	9	AF206665
39	692.6	89.8	1206	9	BC028059
40	686.2	89.0	1154	6	AR080459
41	686.2	89.0	1154	9	HUMTRY
42	630.2	81.7	768	9	AF206666
43	597.6	77.5	918	9	AF206667
44	539.4	70.0	661	9	AF206664
45	519.6	67.4	995	4	DOGMCIRPA

ALIGNMENTS

RESULT 1
AX347853 771 bp DNA linear PAT 01-FEB-2002
LOCUS Sequence 20 from Patent WO0198470.
DEFINITION AX347853
ACCESSION AX347853
VERSION AX347853.1 GI:18495660
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Mafit,M., Niles,A.L. and Haak-Frendscho,M.
TITLE Recombinant proteolytic tryptases, active site mutants thereof, and
methods of making same

JOURNAL Patent: WO 0198470-A 20 27-DEC-2001;
PROMEGA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..771
/organism="Homo sapiens"
/db_xref="taxon:9606"
<1..756
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD2420.1"
/db_xref="GI:1849561"
/translation="GPLEKRIVGQEPARSKMPQVSLRVHGPYMHFGGSLIHPOV
VLTPAACGVDPVKDLALRYQLREOHLVYQDOLPYRIIVHPOFYTAQIGADLALLE
LEEVKVSXVHVTLLPASETEPFGMPGVTGVDNDELRPPFPLKQVPIEMH
NHICDAKYHLGAVTGDDVRIYRDMLCAGNTRRDSQSGSGLVCKVNGTWLQAGVY
SMEGCAQPNRPGITRYTVRYLLDMIHVYRKP"
BASE COUNT 147 a 256 c 244 g 124 t
ORIGIN
Query Match 100.0%; Score 771; DB 6; Length 771;
Best Local Similarity 100.0%; Pred. No. 1.4e-117;
Matches 771; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGCCCTCGAGAAAAGATGTCGGGGGTGAGAGGCCCCAGAGCAAGTGGCCCTGG 60
DB 1 GGGCCCTCGAGAAAAGATGTCGGGGGTGAGAGGCCCCAGAGCAAGTGGCCCTGG 60
QY 61 CAGGTGACCTCGAGAGTCCAGGCCCCATCTGATGATGCACTTCTGGGGGCTCCCTCATC 120
DB 61 CAGGTGACCTCGAGAGTCCAGGCCCCATCTGATGATGCACTTCTGGGGGCTCCCTCATC 120
QY 121 CACCCCAAGTGGGTGACCGCGCGCGCTGCGTGGGAGCGAGAGTCAAGGATCTGGCC 180
DB 121 CACCCCAAGTGGGTGACCGCGCGCGCTGCGTGGGAGCGAGAGTCAAGGATCTGGCC 180
QY 181 GCCCTCAGGGTGAACCTGCGGAGAGCACTCTACTACAGAGCAAGCTGCTGCCGCTC 240
DB 181 GCCCTCAGGGTGAACCTGCGGAGAGCACTCTACTACAGAGCAAGCTGCTGCCGCTC 240
QY 241 AGCAGGATCATCTGTCACCCAGTCTTACACCGCCGAGTCCGAGGAGACATGCCCTTG 300
DB 241 AGCAGGATCATCTGTCACCCAGTCTTACACCGCCGAGTCCGAGGAGACATGCCCTTG 300
QY 301 CTGGAGCTGAGAGACCGGTGAAGGTCTCCAGCCAGTCCACACGGTCAACCTGGCCCT 360
DB 301 CTGGAGCTGAGAGACCGGTGAAGGTCTCCAGCCAGTCCACACGGTCAACCTGGCCCT 360
QY 361 GCCTCAGAGACCTTCCCCCGGGGATGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 361 GCCTCAGAGACCTTCCCCCGGGGATGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 421 AATGATGAGCGCCCTCCACCGCCATTTCTCTGAAGCAAGTGAAGTCCCATATGGA 480
DB 421 AATGATGAGCGCCCTCCACCGCCATTTCTCTGAAGCAAGTGAAGTCCCATATGGA 480
QY 481 AACCAATTTTGAAGCAAAATACACCTTGGCGCTACAGGGAGAGAGAGCTCCGCA 540
DB 481 AACCAATTTTGAAGCAAAATACACCTTGGCGCTACAGGGAGAGAGAGCTCCGCA 540
QY 541 GTCCGTGAGCAATGCTGTGTGCGCGGAACACCGGAGGAGACTATGCCAGGCGCA 600
DB 541 GTCCGTGAGCAATGCTGTGTGCGCGGAACACCGGAGGAGACTATGCCAGGCGCA 600
QY 601 GGAGGGCCCCCTGGTGTGAAGTGAATGGACATCTGCTGCAAGGCGGGGTGGTGA 660
DB 601 GGAGGGCCCCCTGGTGTGAAGTGAATGGACATCTGCTGCAAGGCGGGGTGGTGA 660
QY 661 GGCGAGGCTGTGCGCCAGCCCAACCGGCTGGCATCTACACCGGTGCTACTACTTG 720
DB 661 GGCGAGGCTGTGCGCCAGCCCAACCGGCTGGCATCTACACCGGTGCTACTACTTG 720
QY 721 GACTGGATCCACCACTATGTCCCAAAAAAGCGGTGAAGCGGCGCGCTGCT 771
DB 721 GACTGGATCCACCACTATGTCCCAAAAAAGCGGTGAAGCGGCGCGCTGCT 771

DB 721 GACTGGATCCACCACTATGTCCCAAAAAAGCGGTGAAGCGGCGCGCTGCT 771
RESULT 2
AX347869
LOCUS AX347869 771 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 36 from Patent WO0198470.
ACCESSION AX347869
VERSION AX347869.1 GI:18495676
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Maffei, M., Niles, A. L. and Haak-Frendscho, M.
TITLE Methods of making same
JOURNAL Methods of making same
PROMEGA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..771
/organism="Homo sapiens"
/db_xref="taxon:9606"
<7..756
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD2428.1"
/db_xref="GI:18495677"
/translation="LEKRIVGQEPARSKMPQVSLRVHGPYMHFGGSLIHPOV
VLTPAACGVDPVKDLALRYQLREOHLVYQDOLPYRIIVHPOFYTAQIGADLALLE
LEEVKVSXVHVTLLPASETEPFGMPGVTGVDNDELRPPFPLKQVPIEMH
ICDAKYHLGAVTGDDVRIYRDMLCAGNTRRDSQSGSGLVCKVNGTWLQAGVY
SMEGCAQPNRPGITRYTVRYLLDMIHVYRKP"
BASE COUNT 147 a 257 c 243 g 124 t
ORIGIN
Query Match 99.8%; Score 769.4; DB 6; Length 771;
Best Local Similarity 99.9%; Pred. No. 2.5e-117;
Matches 770; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGCCCTCGAGAAAAGATGTCGGGGGTGAGAGGCCCCAGAGCAAGTGGCCCTGG 60
DB 1 GGGCCCTCGAGAAAAGATGTCGGGGGTGAGAGGCCCCAGAGCAAGTGGCCCTGG 60
QY 61 CAGGTGACCTCGAGAGTCCAGGCCCCATCTGATGATGCACTTCTGGGGGCTCCCTCATC 120
DB 61 CAGGTGACCTCGAGAGTCCAGGCCCCATCTGATGATGCACTTCTGGGGGCTCCCTCATC 120
QY 121 CACCCCAAGTGGGTGACCGCGCGCGCTGCGTGGGAGCGAGAGTCAAGGATCTGGCC 180
DB 121 CACCCCAAGTGGGTGACCGCGCGCGCTGCGTGGGAGCGAGAGTCAAGGATCTGGCC 180
QY 181 GCCCTCAGGGTGAACCTGCGGAGAGCACTCTACTACAGAGCAAGCTGCTGCCGCTC 240
DB 181 GCCCTCAGGGTGAACCTGCGGAGAGCACTCTACTACAGAGCAAGCTGCTGCCGCTC 240
QY 241 AGCAGGATCATCTGTCACCCAGTCTTACACCGCCGAGATCGGAGGGAATGCGCCTG 300
DB 241 AGCAGGATCATCTGTCACCCAGTCTTACACCGCCGAGATCGGAGGGAATGCGCCTG 300
QY 301 CTGGAGCTGAGAGACCGGTGAAGGTCTCCAGCCAGTCCACACGGTCAACCTGGCCCT 360
DB 301 CTGGAGCTGAGAGACCGGTGAAGGTCTCCAGCCAGTCCACACGGTCAACCTGGCCCT 360
QY 361 GCCTCAGAGACCTTCCCCCGGGGATGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 361 GCCTCAGAGACCTTCCCCCGGGGATGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 421 AATGATGAGCGCCCTCCACCGCCATTTCTCTGAAGCAAGTGAAGTCCCATATGGA 480
DB 421 AATGATGAGCGCCCTCCACCGCCATTTCTCTGAAGCAAGTGAAGTCCCATATGGA 480

QY 481 AACCAATTTGTGAGCAAAATACCACTTGGCCCTACACGGAGAGACAGCTCCGATC 540
|||||
Db 481 AACCAATTTGTGAGCAAAATACCACTTGGCCCTACACGGAGAGAGAGCTCCGATC 540
QY 541 GTCCGTGACGACATCTGTGTGCGGGAAACACCCGGAGGACTCATGCCAGGCGCATCC 600
|||||
Db 541 GTCCGTGACGACATCTGTGTGCGGGAAACACCCGGAGGACTCATGCCAGGCGCATCC 600
QY 601 GGAGGGCCCTGTGTGCAAGTGAATGGACACTGGCTGCAGAGCGGGGTGGTCACTGG 660
|||||
Db 601 GGAGGGCCCTGTGTGCAAGTGAATGGACACTGGCTGCAGAGCGGGGTGGTCACTGG 660
QY 661 GGCGAGGGCTGTGTGCGCCAGCCCAACCGGCTGTGCATCTACACCCGTGTCACTACTTGG 720
|||||
Db 661 GGCGAGGGCTGTGTGCGCCAGCCCAACCGGCTGTGCATCTACACCCGTGTCACTACTTGG 720
QY 721 GACTGGATTCACCACTATGTCTCCCAAAAAAGCCGTGAAGCGGCGCGCTGT 771
|||||
Db 721 GACTGGATTCACCACTATGTCTCCCAAAAAAGCCGTGAAGCGGCGCGCTGT 771

RESULT 3

LOCUS AR165112 771 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 4 from patent US 6274366.
ACCESSION AR165112
VERSION AR165112.1 GI:16238527
KEYWORDS

SOURCE

ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 771)
Maffitt,M.A., Niles,A.L. and Haak-Frendscho,M.

TITLE Enzymatically-active recombinant human .beta.-tryptase and method
of making same

JOURNAL Patent: US 6274366-A 4 14-AUG-2001;
FEATURES Location/Qualifiers

source

BASE COUNT 149 a 256 c 242 g 124 t
ORIGIN

Query Match 99.2%; Score 764.6; DB 6; Length 771;
Best Local Similarity 99.5%; Pred. No. 1.6e-116;
Matches 767; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGGCCCCCTCGAAGAAAGATCTGCGGGGTCAAGAGGCCCCCAGAGCAAGTGGCCCTGG 60
|||||
Db 1 GGGCCCCCTCGAAGAAAGATCTGCGGGGTCAAGAGGCCCCCAGAGCAAGTGGCCCTGG 60
QY 61 CAGGTGAGCTTGAGAGTCCAGCGGCCCATCTACTGATGACACTTCTGCGGGGCTCCCTCATC 120
|||||
Db 61 CAGGTGAGCTTGAGAGTCCAGCGGCCCATCTACTGATGACACTTCTGCGGGGCTCCCTCATC 120
QY 121 CACCCCAAGTGGGTCTGACCGCCGCGCTGCGGGAGCCGAGAGTCAAGAGTCTGGCC 180
|||||
Db 121 CACCCCAAGTGGGTCTGACCGCCGCGCTGCGGGAGCCGAGAGTCAAGAGTCTGGCC 180
QY 181 GCGCTCAGAGGTGCACTGCGGGAGAGCAACCTCTACTACAGAGCAGCTGCTGCCGGTC 240
|||||
Db 181 GCGCTCAGAGGTGCACTGCGGGAGAGCAACCTCTACTACAGAGCAGCTGCTGCCGGTC 240
QY 241 AGCAGGATCATGTGTCACGACAGTTCTACACCGCCCGCATGGAGCGACATCGCCCTG 300
|||||
Db 241 AGCAGGATCATGTGTCACGACAGTTCTACACCGCCCGCATGGAGCGACATCGCCCTG 300
QY 301 CTGAGAGCTGAGAGACCGGTGAAGTCTTCACAGCAAGTCCACACGCTCACTTGGCCCT 360
|||||
Db 301 CTGAGAGCTGAGAGACCGGTGAAGTCTTCACAGCAAGTCCACACGCTCACTTGGCCCT 360
QY 361 GCGCTCAGAGAGCTTCCCGCGGGAGATCGCGTCTGGGTCACTGGCTGGGCGATGTGAC 420
|||||
Db 361 GCGCTCAGAGAGCTTCCCGCGGGAGATCGCGTCTGGGTCACTGGCTGGGCGATGTGAC 420

QY 421 AATGATGAGCGGCTCCCAACCGCATTTCTCTGAAGCAGGTGAAGTCCCATTAATGGA 480
|||||
Db 421 AATGATGAGCGGCTCCCAACCGCATTTCTCTGAAGCAGGTGAAGTCCCATTAATGGA 480
QY 481 AACCAATTTGTGAGCAAAATACCACTTGGCGGCTACACGGAGAGACAGCTCCGATC 540
|||||
Db 481 AACCAATTTGTGAGCAAAATACCACTTGGCGGCTACACGGAGAGAGAGCTCCGATC 540
QY 541 GTCCGTGACGACATCTGTGTGCGGGAAACACCCGGAGGACTCATGCCAGGCGCATCC 600
|||||
Db 541 GTCCGTGACGACATCTGTGTGCGGGAAACACCCGGAGGACTCATGCCAGGCGCATCC 600
QY 601 GGAGGGCCCTGTGTGCAAGTGAATGGACACTGGCTGCAGAGCGGGGTGGTCACTGG 660
|||||
Db 601 GGAGGGCCCTGTGTGCAAGTGAATGGACACTGGCTGCAGAGCGGGGTGGTCACTGG 660
QY 661 GGCGAGGGCTGTGTGCGCCAGCCCAACCGGCTGTGCATCTACACCCGTGTCACTACTTGG 720
|||||
Db 661 GGCGAGGGCTGTGTGCGCCAGCCCAACCGGCTGTGCATCTACACCCGTGTCACTACTTGG 720
QY 721 GACTGGATTCACCACTATGTCTCCCAAAAAAGCCGTGAAGCGGCGCGCTGT 771
|||||
Db 721 GACTGGATTCACCACTATGTCTCCCAAAAAAGCCGTGAAGCGGCGCGCTGT 771

RESULT 4

LOCUS AX347841 771 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 8 from Patent WO0198470.
ACCESSION AX347841
VERSION AX347841.1 GI:18495648
KEYWORDS

SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1
Maffitt,M., Niles,A.L. and Haak-Frendscho,M.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

TITLE Recombinant proteolytic trypsinases, active site mutants thereof, and
methods of making same

JOURNAL Patent: WO 0198470-A 8 27-DEC-2001;
FEATURES PROMEGA CORPORATION (US)

source Location/Qualifiers

1..771

/organism="Homo sapiens"

/db_xref="taxon:9606"

<7..756

/note="unnamed protein product"

/codon_start=1

/protein_id="CAD22418.1"

/db_xref="GI:18495648"

/translation="LEKRIYGGQAPSKWFWVSLRVHGFYWMHFCGSLIHPOWVL
TAAHCVGPVDVLAALRYOLREOHLYVODLIPVSRILVHPOFYTAIGADIALIELE
EPVAVSHVHTVTLPPASEPTEPMPGMCVWGTGMDVNDERLPPFPKQVYPIEMENH
ICDKYHNGATYGDVDRIVRDMLCAGTRRDSQDGSGLVCKVNGTMIDQAVYSW
GEGCAOPNRPIYRVRYYLIDWHIHYPKR"

BASE COUNT 149 a 256 c 242 g 124 t

ORIGIN

Query Match 99.2%; Score 764.6; DB 6; Length 771;
Best Local Similarity 99.5%; Pred. No. 1.6e-116;
Matches 767; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGGCCCCCTCGAAGAAAGATCTGCGGGGTCAAGAGGCCCCCAGAGCAAGTGGCCCTGG 60
|||||
Db 1 GGGCCCCCTCGAAGAAAGATCTGCGGGGTCAAGAGGCCCCCAGAGCAAGTGGCCCTGG 60
QY 61 CAGGTGAGCTTGAGAGTCCAGCGGCCCATCTACTGATGACACTTCTGCGGGGCTCCCTCATC 120
|||||
Db 61 CAGGTGAGCTTGAGAGTCCAGCGGCCCATCTACTGATGACACTTCTGCGGGGCTCCCTCATC 120
QY 121 CACCCCAAGTGGGTCTGACCGCCGCGCTGCGGGAGCCGAGAGTCAAGAGTCTGGCC 180
|||||
Db 121 CACCCCAAGTGGGTCTGACCGCCGCGCTGCGGGAGCCGAGAGTCAAGAGTCTGGCC 180

```
|||||
Db 121 CACCCCAAGTGGTGTGACCGGACGACCTGCTGGAGCCGGACGTCAGAGATCTGGCC 180
Qy 181 GCCCTCAGGGGCGAAGTCGGGGAGCAGACCTCTACTACAGAGACGACGTCCTCCGGTC 240
Db 181 GCCCTCAGGGGCGAAGTCGGGGAGCAGACCTCTACTACAGAGACGACGTCCTCCGGTC 240
Qy 241 AGCAGATCATCTGTGACCCACAGTTCCTACACGCCCCAGATCGAGCGGAGCATCGCCCTG 300
Db 241 AGCAGATCATCTGTGACCCACAGTTCCTACACGCCCCAGATCGAGCGGAGCATCGCCCTG 300
Qy 301 CTGAGAGTCGAGAGCGCGGTGAAGTCTCCAGCCAGCTCCACAGGTCACCTCTCCCTT 360
Db 301 CTGAGAGTCGAGAGCGCGGTGAAGTCTCCAGCCAGCTCCACAGGTCACCTCTCCCTT 360
Qy 361 GCCTCAGAGACCTTCCCGCGGGATGCGTGGGTCATCTGGCTGGGGGAGATGTGGAC 420
Db 361 GCCTCAGAGACCTTCCCGCGGGATGCGTGGGTCATCTGGCTGGGGGAGATGTGGAC 420
Qy 421 AATGATGAGCCCTCCACCCCAATTCCTCTGAAGCAGGTGAAGTCCCATTAATGAA 480
Db 421 AATGATGAGCCCTCCACCCCAATTCCTCTGAAGCAGGTGAAGTCCCATTAATGAA 480
Qy 481 AACCAATTTGTGACGCAAAATACACCTTGGCGCTACAGGGAGACAGATCGCGCATC 540
Db 481 AACCAATTTGTGACGCAAAATACACCTTGGCGCTACAGGGAGAGATCGCGCATC 540
Qy 541 GTCCGTGAGCAGATGCTGTGCGCGGGAACACCCGAGGAGATCTACAGGGGAGATC 600
Db 541 GTCCGTGAGCAGATGCTGTGCGCGGGAACACCCGAGGAGATCTACAGGGGAGATC 600
Qy 601 GGAGGGCCCTGTGTGTGCAAGTGAATGGACCTGTGTCAGAGCGCGGTGGTCACTGG 660
Db 601 GGAGGGCCCTGTGTGTGCAAGTGAATGGACCTGTGTCAGAGCGCGGTGGTCACTGG 660
Qy 661 GGCGAGGGCTGTGCGGACCCCAACCGGCTGGCATCTACACCGGCTGCTACTACTTG 720
Db 661 GGCGAGGGCTGTGCGGACCCCAACCGGCTGGCATCTACACCGGCTGCTACTACTTG 720
Qy 721 GACTGGATCCACACTATGTGCCCCAAAAAGCGGTGAAGCGCGCGCTGCT 771
Db 721 GACTGGATCCACACTATGTGCCCCAAAAAGCGGTGAAGCGCGCGCTGCT 771

RESULT 5
AX347838 771 bp DNA linear PAT 01-FEB-2002
LOCUS AX347838
DEFINITION Sequence 5 from Patent WO0198470.
ACCESSION AX347838
VERSION AX347838.1 GI:18495645
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Maffei,M., Niles,A.L. and Haak-Frendscho,M.
AUTHORS Recombinant proteolytic tryptases, active site mutants thereof, and
TITLE Methods of making same
JOURNAL Patent: WO 0198470-A 5 27-DEC-2001;
PROMEGA CORPORATION (US)
FEATURES
SOURCE Location/Qualifiers
1..771
/organism="Homo sapiens"
/db_xref="taxon:9606"
<7..756
/note="unnamed protein product"
CDS
/codon_start=1
/protein_id="CAD22417.1"
/db_xref="GI:18495646"
/translation="LEKRIYVGQEAPEPSKWPMOVSLRVHGPYMMHFCGSLIHPOWVL
TRAHCYGDVYDLALRYOLRPOHLXODOLIPYRITIVHPOFTYAGADIALLELE
EPVNVSSHVHTVTLPPASETPPGMPCWVTGMDVNDERLPPPEPLAQVAVPIWENH
```

```
ICDAXHGLSAYTGDVYRIVRDMILCAGNTRRPSQGGSGPLVCKVNGTWLQAGVSW
GEGCADPNRPIYTRVYTYLDMTHNHYPRKP"
BASE COUNT 149 a 257 c 241 g 124 t
ORIGIN
Query Match 99.0%; Score 763; DB 6; Length 771;
Best Local Similarity 99.4%; Pred. No. 2,9e-116;
Matches 766; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 GGGCCCTCGAAGAAAGATGTCGGGGGTGAGAGGCCCCCAGAGCAAGTGGCCCTG 60
Db 1 GGGCCCTCGAAGAAAGATGTCGGGGGTGAGAGGCCCCCAGAGCAAGTGGCCCTG 60
Qy 61 CAGGTGAGCCTGAGAGTCCACGGCCATCTGATGATGACATTTCTGGGGGCTCCCTATC 120
Db 61 CAGGTGAGCCTGAGAGTCCACGGCCATCTGATGATGACATTTCTGGGGGCTCCCTATC 120
Qy 121 CACCCCAAGTGGTGTGACCGCGCGGCTGCTGGGAGCGGACGTCAAGAGATCTGGC 180
Db 121 CACCCCAAGTGGTGTGACCGCGCGGCTGCTGGGAGCGGACGTCAAGAGATCTGGC 180
Qy 181 GCCCTCAGGGTGTCAACTGCGGGAGCAGACCTCTACTACAGAGACAGTCTGCGGTC 240
Db 181 GCCCTCAGGGTGTCAACTGCGGGAGCAGACCTCTACTACAGAGACAGTCTGCGGTC 240
Qy 241 AGCAGATCATCTGTGACCCACAGTTCCTACACGCCCCAGATCGGAGGGAGCATCGCCCTG 300
Db 241 AGCAGATCATCTGTGACCCACAGTTCCTACACGCCCCAGATCGGAGGGAGCATCGCCCTG 300
Qy 301 CTGAGAGTCGAGAGCGCGGTGAAGTCTCCAGCAGCTCCACAGGTCACCTCCCTCCT 360
Db 301 CTGAGAGTCGAGAGCGCGGTGAAGTCTCCAGCAGCTCCACAGGTCACCTCCCTCCT 360
Qy 361 GCCTCAGAGACCTTCCCGCGGGATGCGTGGGTCATCTGGCTGGGGGAGATGTGGAC 420
Db 361 GCCTCAGAGACCTTCCCGCGGGATGCGTGGGTCATCTGGCTGGGGGAGATGTGGAC 420
Qy 421 AATGATGAGCCCTCCACCCCAATTCCTCTGAAGCAGGTGAAGTCCCATTAATGAA 480
Db 421 AATGATGAGCCCTCCACCCCAATTCCTCTGAAGCAGGTGAAGTCCCATTAATGAA 480
Qy 481 AACCAATTTGTGACGCAAAATACACCTTGGCGCTACAGGGAGACAGATCGCGCATC 540
Db 481 AACCAATTTGTGACGCAAAATACACCTTGGCGCTACAGGGAGACAGATCGCGCATC 540
Qy 541 GTCCGTGAGCAGATGCTGTGCGCGGAACACCCGAGGAGATCTACAGGGGAGATC 600
Db 541 GTCCGTGAGCAGATGCTGTGCGCGGAACACCCGAGGAGATCTACAGGGGAGATC 600
Qy 601 GGAGGGCCCTGTGTGTGCAAGTGAATGGACCTGTGTCAGAGCGCGGTGGTCACTGG 660
Db 601 GGAGGGCCCTGTGTGTGCAAGTGAATGGACCTGTGTCAGAGCGCGGTGGTCACTGG 660
Qy 661 GGCGAGGGCTGTGCGGACCCCAACCGGCTGGCATCTACACCGGCTGCTACTACTTG 720
Db 661 GGCGAGGGCTGTGCGGACCCCAACCGGCTGGCATCTACACCGGCTGCTACTACTTG 720
Qy 721 GACTGGATCCACACTATGTGCCCCAAAAAGCGGTGAAGCGCGCGCTGCT 771
Db 721 GACTGGATCCACACTATGTGCCCCAAAAAGCGGTGAAGCGCGCGCTGCT 771

RESULT 6
AX347855 771 bp DNA linear PAT 01-FEB-2002
LOCUS AX347855
DEFINITION Sequence 22 from Patent WO0198470.
ACCESSION AX347855
VERSION AX347855.1 GI:18495662
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
```


REFERENCE 1
 AUTHORS Maffei, M., Niles, A. L., and Haak-Frendscho, M.
 TITLE Recombinant proteolytic trypsinases, active site mutants thereof, and methods of making same
 JOURNAL Patent: WO 0198470-A 22 27-DEC-2001;
 PROMEGA CORPORATION (US)
 FEATURES location/Qualifiers
 source 1..771
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 <7..756
 CDS /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD22421.1"
 /db_xref="GI:18495678"
 /translation="LEKRIIVGGQEPNRSKIPWQVSLRVHGPYMHFGGSLIHQWVL
 TAAHCVGPDAVDLALRVQLREQHLYYODQLPVSRILVHPQFYTAQGAIALLELE
 EPVNVSSHVHTVTLPPASEIFPPGPMGWTDGVDNDRLEPPEFLKQVAVPIEMEN
 ICDAKYHLGATGDVRIYRDMDCAGSTRDSCQDGSGLVCKVNGTWTLOAGVVS
 GEGCAOPRPIRYIRVYLLDMHIVYKRP"
 BASE COUNT 149 a 256 c 242 g 124 t
 ORIGIN
 Query Match 98.8%; Score 761.4; DB 6; Length 771;
 Best Local Similarity 99.2%; Pred. No. 5.3e-116;
 Matches 765; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 661 GGCAGAGGCTGTGCCAGGCCAACCGGCTGAGCANTACACCCGTGTCACCTACTACTTG 720
 QY 721 GAGTGGATCCACCTAGTCCCAAAAAAGCGTGAAGGGCGCGCTGT 771
 Db 721 GAGTGGATCCACCTAGTCCCAAAAAAGCGTGAAGGGCGCGCTGT 771
 RESULT 7
 AX347871
 LOCUS AX347871 771 bp DNA linear PAT 01-FEB-2002
 DEFINITION Sequence 38 from Patent WO0198470.
 ACCESSION AX347871
 VERSION AX347871.1 GI:18495678
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Maffei, M., Niles, A. L., and Haak-Frendscho, M.
 TITLE Recombinant proteolytic trypsinases, active site mutants thereof, and methods of making same
 JOURNAL Patent: WO 0198470-A 38 27-DEC-2001;
 PROMEGA CORPORATION (US)
 FEATURES location/Qualifiers
 source 1..771
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 <7..756
 CDS /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD22429.1"
 /db_xref="GI:18495679"
 /translation="LEKRIIVGGQEPNRSKIPWQVSLRVHGPYMHFGGSLIHQWVL
 TAAHCVGPDAVDLALRVQLREQHLYYODQLPVSRILVHPQFYTAQGAIALLELE
 EPVNVSSHVHTVTLPPASEIFPPGPMGWTDGVDNDRLEPPEFLKQVAVPIEMEN
 ICDAKYHLGATGDVRIYRDMDCAGSTRDSCQDGSGLVCKVNGTWTLOAGVVS
 GEGCAOPRPIRYIRVYLLDMHIVYKRP"
 BASE COUNT 149 a 257 c 241 g 124 t
 ORIGIN
 Query Match 98.5%; Score 759.8; DB 6; Length 771;
 Best Local Similarity 99.1%; Pred. No. 9.7e-116;
 Matches 764; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

BASE COUNT 152 a 255 c 241 g 123 t
ORIGIN

GGCGAOPNREIYTRVYYLDWIIHHVPRKP"

Query Match 97.9%; Score 755; DB 6; Length 771;
Best Local Similarity 98.7%; Pred. No. 6e-115;
Matches 761; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 GGGCCCCCTCGAGAAAAAGATCGTGGGGGTCAGAGAGCCCCCAGAGAAAGTGGCCCTGG 60
DB 1 GGGCCCCCTCGAGAAAAAGATCGTGGGGGTCAGAGAGCCCCCAGAGAAAGTGGCCCTGG 60
OY 61 CAGGTGAGCCTGAGAGTCCAGGCCCATCTGATGCACTTCTGCGGGGCTCCCTCATC 120
DB 61 CAGGTGAGCCTGAGAGTCCAGGCCCATCTGATGCACTTCTGCGGGGCTCCCTCATC 120
OY 121 CACCCCCAGTGGTGTCTACCCCGCGCGCTGCGTGGGACCGGACGCTCAAGATCTGGCC 180
DB 121 CACCCCCAGTGGTGTCTACCCCGCGCGCTGCGTGGGACCGGACGCTCAAGATCTGGCC 180
OY 181 GCGCTCAGAGGTCACACTGCGGGAGGAGCAGCTCTACTACAGAGCAGCTGCTGCCGGTC 240
DB 181 GCGCTCAGAGGTCACACTGCGGGAGGAGCAGCTCTACTACAGAGCAGCTGCTGCCGGTC 240
OY 241 AGCAGGATCATCTGTCACCCAGCTTCTACACCGCCAGATCGAGCGACATCGCCCTG 300
DB 241 AGCAGGATCATCTGTCACCCAGCTTCTACACCGCCAGATCGAGCGACATCGCCCTG 300
OY 301 CTGGAGCTGAGAGACCGGTGAAGTCTTCACGCCACGTCACACGCTCAACCTGCCCCCT 360
DB 301 CTGGAGCTGAGAGACCGGTGAAGTCTTCACGCCACGTCACACGCTCAACCTGCCCCCT 360
OY 361 GCGCTCAGAGGTCACACTGCGGGAGGAGCAGCTCTACTACAGAGCAGCTGCTGCCGGTC 420
DB 361 GCGCTCAGAGGTCACACTGCGGGAGGAGCAGCTCTACTACAGAGCAGCTGCTGCCGGTC 420
OY 421 AATGATGAGCGCGCTCCCAACCGCATTTCTGGAAGCAGGTGAAGTCCCATTAATGAA 480
DB 421 AATGATGAGCGCGCTCCCAACCGCATTTCTGGAAGCAGGTGAAGTCCCATTAATGAA 480
OY 481 AACCAATTTTGTGAGAGCAAAATACCACTTGGCGCTACAGGGAGAGAGCAGCTCCGCATC 540
DB 481 AACCAATTTTGTGAGAGCAAAATACCACTTGGCGCTACAGGGAGAGAGCAGCTCCGCATC 540
OY 541 GTCCCTGAGAGATCTGTGTGCGGGGAACACCGAGGGAGTCAATGACAGGGGAGCATCC 600
DB 541 GTCCCTGAGAGATCTGTGTGCGGGGAACACCGAGGGAGTCAATGACAGGGGAGCATCC 600
OY 601 GGAGGGCCCCGTGGTGAAGGTGAATGACACTGGCTGACAGCGGGGCTGTGTCAGCTGG 660
DB 601 GGAGGGCCCCGTGGTGAAGGTGAATGACACTGGCTGACAGCGGGGCTGTGTCAGCTGG 660
OY 661 GCGGAGGGCTGTGCGGACGCCACCGGCTGACATCTACACCGGCTGCTCACTACTTGG 720
DB 661 GCGGAGGGCTGTGCGGACGCCACCGGCTGACATCTACACCGGCTGCTCACTACTTGG 720
OY 721 GACTGATTCACACACTATGTGCCCCAAAAAGCCGTGAAGCGGCGCGCTGT 771
DB 721 GACTGATTCACACACTATGTGCCCCAAAAAGCCGTGAAGCGGCGCGCTGT 771

RESULT 10
LOCUS AX347873 771 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 40 from Patent WO0198470.
ACCESSION AX347873
VERSION AX347873.1 GI:18495680
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1

AUTHORS Maffei, M., Niles, A. L. and Haak-Frendscho, M.
TITLE Recombinant proteolytic tryptases, active site mutants thereof, and methods of making same
JOURNAL Patent: WO 0198470-A 40 27-DEC-2001;
PROMEGA CORPORATION (US)
FEATURES Location/Qualifiers
source 1..771
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS <1..756
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD22430.1"
/db_xref="GI:18495681"
/translation="GPLEKRYVGDEAPRSMKPMWQVSLRVHGYPWMHFGGSLIHPQW
VLTARHCYPRVKTALALRYDLRQHLIYDQDLIPVRIYHPPTAOLGADLALLE
LEEPYVNSSHVHTVLPPEASITFPPGMCWCTGVDNDERLPPPLKQVPIVE
NHICDARYHLGAYTGDDVRIYRDMMLAGNTRRDSQGDAGGPLVCVNGTWLQAGVY
SMGEGCAOPNRPFIYTRVYYLDWIIHHVPRKP"

BASE COUNT 150 a 256 c 240 g 125 t
ORIGIN

Query Match 97.7%; Score 753.4; DB 6; Length 771;
Best Local Similarity 98.6%; Pred. No. 1.1e-114;
Matches 760; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 1 GGGCCCCCTCGAGAAAAAGATCGTGGGGGTCAGAGAGCCCCCAGAGAAAGTGGCCCTGG 60
DB 1 GGGCCCCCTCGAGAAAAAGATCGTGGGGGTCAGAGAGCCCCCAGAGAAAGTGGCCCTGG 60
OY 61 CAGGTGAGCCTGAGAGTCCAGGCCCATCTGATGCACTTCTGCGGGGCTCCCTCATC 120
DB 61 CAGGTGAGCCTGAGAGTCCAGGCCCATCTGATGCACTTCTGCGGGGCTCCCTCATC 120
OY 121 CACCCCCAGTGGTGTCTACCCCGCGCGCTGCGTGGGACCGGACGCTCAAGATCTGGCC 180
DB 121 CACCCCCAGTGGTGTCTACCCCGCGCGCTGCGTGGGACCGGACGCTCAAGATCTGGCC 180
OY 181 GCGCTCAGAGGTCACACTGCGGGAGGAGCAGCTCTACTACAGAGCAGCTGCTGCCGGTC 240
DB 181 GCGCTCAGAGGTCACACTGCGGGAGGAGCAGCTCTACTACAGAGCAGCTGCTGCCGGTC 240
OY 241 AGCAGGATCATCTGTCACCCAGCTTCTACACCGCCAGATCGAGCGACATCGCCCTG 300
DB 241 AGCAGGATCATCTGTCACCCAGCTTCTACACCGCCAGATCGAGCGACATCGCCCTG 300
OY 301 CTGGAGCTGAGAGACCGGTGAAGTCTTCACGCCACGTCACACGCTCAACCTGCCCCCT 360
DB 301 CTGGAGCTGAGAGACCGGTGAAGTCTTCACGCCACGTCACACGCTCAACCTGCCCCCT 360
OY 361 GCGCTCAGAGGTCACACTGCGGGAGGAGTGGCTGTGTCAGCTGGGCGGATGTGAA 420
DB 361 GCGCTCAGAGGTCACACTGCGGGAGGAGTGGCTGTGTCAGCTGGGCGGATGTGAA 420
OY 421 AATGATGAGCGCGCTCCCAACCGCATTTCTGGAAGCAGGTGAAGTCCCATTAATGAA 480
DB 421 AATGATGAGCGCGCTCCCAACCGCATTTCTGGAAGCAGGTGAAGTCCCATTAATGAA 480
OY 481 AACCAATTTTGTGAGAGCAAAATACCACTTGGCGCTACAGGGAGAGAGCAGCTCCGCATC 540
DB 481 AACCAATTTTGTGAGAGCAAAATACCACTTGGCGCTACAGGGAGAGAGCAGCTCCGCATC 540
OY 541 GTCCCTGAGAGATCTGTGTGCGGGGAACACCGAGGGAGTCAATGACAGGGGAGCATCC 600
DB 541 GTCCCTGAGAGATCTGTGTGCGGGGAACACCGAGGGAGTCAATGACAGGGGAGCATCC 600
OY 601 GGAGGGCCCCGTGGTGAAGGTGAATGACACTGGCTGACAGCGGGGCTGTGTCAGCTGG 660
DB 601 GGAGGGCCCCGTGGTGAAGGTGAATGACACTGGCTGACAGCGGGGCTGTGTCAGCTGG 660
OY 661 GCGGAGGGCTGTGCGGACGCCACCGGCTGACATCTACACCGGCTGCTCACTACTTGG 720
DB 661 GCGGAGGGCTGTGCGGACGCCACCGGCTGACATCTACACCGGCTGCTCACTACTTGG 720

```
OY 721 GACTGATCCACCCTATGTGCCCCAAAAGCCGTGAAGCGCCCGCTCGT 771
|||||
Db 721 GACTGATCCACCCTATGTGCCCCAAAAGCCGTGAAGCGCCCGCTCGT 771

RESULT 11
AX347875 771 bp DNA linear PAT 01-FEB-2002
LOCUS Sequence 42 from Patent WO0198470.
ACCESSION AX347875
VERSION AX347875.1 GI:18495682
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Mafilt,M., Niles,A.L. and Haak-Frendscho,M.
AUTHORS Recombinant proteolytic trypases, active site mutants thereof, and
TITLE Patent: WO 0198470-A 42 27-DEC-2001.
JOURNAL PROMEGA CORPORATION (US)
FEATURES
source Location/Qualifiers
1. 771
/organism="Homo sapiens"
/db_xref="taxon:9606"
<7. 756
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD22431.1"
/translation="LEKRIYGOEAPRSKPMQVSLRVHGRPYMHFGGSLIPQWVL
/TAHGVDPYDLALRVOLREOHLYODLLPVSRITIVRPOYTOIGADILLELE
EPYVSSHVTYVLPASSETPPGMPCWYMGVDNDELRPPPLKQKVPIMENH
ICPAKTHLGAITGDDVRIYRDMILCAGNTRDSCQSDGSPYKVNGLQAGVYSW
GECADQNPRIYRIVTYLDMIHVYPKR"
BASE COUNT 152 a 256 c 240 g 123 t
ORIGIN

Query Match 97.7%; Score 753.4; DB 6; Length 771;
Best Local Similarity 96.6%; Pred. No. 1.1e-114;
Matches 760; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

OY 1 GGGCCCCCGAGAAAAGATGTCGGGGGTGAGAGGCCCCAGAGCAAGTGGCCCTGG 60
|||||
Db 1 GGGCCCCCGAGAAAAGATGTCGGGGGTGAGAGGCCCCAGAGCAAGTGGCCCTGG 60

OY 61 CAGGTGAGCCTGAGAGTCCACGGCCCATCTGATGATGACTTCTCGGGGGCTCCCTCATC 120
|||||
Db 61 CAGGTGAGCCTGAGAGTCCACGGCCCATCTGATGATGACTTCTCGGGGGCTCCCTCATC 120

OY 121 CACCCCAAGTGGTCTGACCGCGCGGCGGTGGGAGCGGAGCTCAAGGATCTGGCC 180
|||||
Db 121 CACCCCAAGTGGTCTGACCGCGCGGCGGTGGGAGCGGAGCTCAAGGATCTGGCC 180

OY 121 CACCCCAAGTGGTCTGACCGCGCGGCGGTGGGAGCGGAGCTCAAGGATCTGGCC 180
|||||
Db 121 CACCCCAAGTGGTCTGACCGCGCGGCGGTGGGAGCGGAGCTCAAGGATCTGGCC 180

OY 181 GCCCTCAGAGTCACTGCGGAGAGCAGCCTCTACTACAGAGCACTGCTGCGGCTC 240
|||||
Db 181 GCCCTCAGAGTCACTGCGGAGAGCAGCCTCTACTACAGAGCACTGCTGCGGCTC 240

OY 241 AGCAGGATCATGTCACCCAGTTCATACACGCCCCAGANTCGAGCGGACATTCGCCCTG 300
|||||
Db 241 AGCAGGATCATGTCACCCAGTTCATACACGCCCCAGANTCGAGCGGACATTCGCCCTG 300

OY 301 CTGGAGCTGGAGAGCCGGTGAAGTCTTCACGCCACGTCACACCGTCAACCTGCCCCCT 360
|||||
Db 301 CTGGAGCTGGAGAGCCGGTGAAGTCTTCACGCCACGTCACACCGTCAACCTGCCCCCT 360

OY 361 GCGTCAGAGACCTTCCCCCGGGGATGCGTGGTCTGCTGCTGCTGCTGCTGCTGCTG 420
|||||
Db 361 GCGTCAGAGACCTTCCCCCGGGGATGCGTGGTCTGCTGCTGCTGCTGCTGCTGCTG 420

OY 421 AATGATGAGCGGCTCCACCGCCATTTCCTCTGAAGCAGGTGAAGTCCCATTAATGAA 480
|||||
```

```
Db 421 AATGATGAGCGGCTCCACCGCCATTTCCTCTGAAGCAGGTGAAGTCCCATTAATGAA 480
|||||
OY 481 AACCATATTGTGACGCAAAATACACACTTGGCGCTTACACGGGAGACAGCTCCGATC 540
|||||
Db 481 AACCATATTGTGACGCAAAATACACACTTGGCGCTTACACGGGAGACAGCTCCGATC 540

OY 541 GTCCGTGACGACATGCTGTGTGTCGGGGGAAACACCCGGAGGACATGACGAGGCACTCC 600
|||||
Db 541 GTCCGTGACGACATGCTGTGTGTCGGGGGAAACACCCGGAGGACATGACGAGGCACTCC 600

OY 601 GGAGGGCCCTGTGTGTCAGAGTGAATGGACACTGCTGAGCGGGGCGTGTCACTGG 660
|||||
Db 601 GGAGGGCCCTGTGTGTCAGAGTGAATGGACACTGCTGAGCGGGGCGTGTCACTGG 660

OY 661 GGGAGGCGCTGTGCCAGCCCAACCGGCTGCACTTACACCGCTGTCACTTACTTG 720
|||||
Db 661 GGGAGGCGCTGTGCCAGCCCAACCGGCTGCACTTACACCGCTGTCACTTACTTG 720

OY 721 GACTGATCCACCCTATGTGCCCCAAAAGCCGTGAAGCGCCCGCTCGT 771
|||||
Db 721 GACTGATCCACCCTATGTGCCCCAAAAGCCGTGAAGCGCCCGCTCGT 771

RESULT 12
AX347861 735 bp DNA linear PAT 01-FEB-2002
LOCUS Sequence 28 from Patent WO0198470.
ACCESSION AX347861
VERSION AX347861.1 GI:18495668
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Mafilt,M., Niles,A.L. and Haak-Frendscho,M.
AUTHORS Recombinant proteolytic trypases, active site mutants thereof, and
TITLE Patent: WO 0198470-A 28 27-DEC-2001.
JOURNAL PROMEGA CORPORATION (US)
FEATURES
source Location/Qualifiers
1. 735
/organism="Homo sapiens"
/db_xref="taxon:9606"
<1. >735
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD22424.1"
/translation="IVGGQFAPRSKPMQVSLRVHGRPYMHFGGSLIPQWVLTA
CVGPDVADLALRVOLREOHLYODLLPVSRITIVRPOYTOIGADILLELEPYK
VSSHVTYVLPASSETPPGMPCWYMGVDNDELRPPPLKQKVPIMENHICDA
KYHLGATGDDVRIYRDMILCAGNTRDSCQSDGSPYKVNGLQAGVYSWGBGC
AQPRIYRIVTYLDMIHVYPKR"
BASE COUNT 139 a 245 c 231 g 120 t
ORIGIN

Query Match 95.3%; Score 735; DB 6; Length 735;
Best Local Similarity 100.0%; Pred. No. 1.2e-111;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 ATCGTGGGGGTGAGAGGCCCCCGAGAGCAAGTGGCCCTGGCAAGTGAAGATC 78
|||||
Db 1 ATCGTGGGGGTGAGAGGCCCCCGAGAGCAAGTGGCCCTGGCAAGTGAAGATC 78

OY 79 CACGGCCCATCTGATGATCATTCTGCGGGGCTCCCATCCACCCAGTGGGTGCTG 138
|||||
Db 79 CACGGCCCATCTGATGATCATTCTGCGGGGCTCCCATCCACCCAGTGGGTGCTG 138

OY 139 ACCGCGGGCGGTGCTGGTGGAGACCGGACGTCAAGATCTGCGCGCCCTCAAGGTTGCAACTG 198
|||||
Db 139 ACCGCGGGCGGTGCTGGTGGAGACCGGACGTCAAGATCTGCGCGCCCTCAAGGTTGCAACTG 198

OY 121 ACCGCGCGGCGGTGCTGGTGGAGACCGGACGTCAAGATCTGCGCGCCCTCAAGGTTGCAACTG 180
|||||
```



```

Db 627 TGTGTGCGGGGAACACCCGGAGGACTCATGCCAGGGCGACTCCGAGGGCCCTGTGTGT 686
OY 617 GCAAGGTGAATGACACCTGCTGCAGCGCGGGCTGTGTCAGTGGGGCGAGGGCTGTGCCC 676
    |||||||
Db 687 GCAAGGTGAATGACACCTGCTGCAGCGCGGGCTGTGTCAGTGGGGCGAGGGCTGTGCCC 746
OY 677 AGCCCAACCGGCTGTGCACTTACACCCGCTGTGACCTACTTACTTGGACTGTGATCCACCACT 736
    |||||||
Db 747 AGCCCAACCGGCTGTGCACTTACACCCGCTGTGACCTACTTACTTGGACTGTGATCCACCACT 806
OY 737 ATGTCCCAAAAAAGCCGTGAAGCGGC 763
    |||||||
Db 807 ATGTCCCAAAAAAGCCGTGAAGTCAGGC 833
    |||||||

```

Search completed: January 31, 2003, 06:12:33
 Job time : 2817 secs

7

100